

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|--------------------|----------------------|
| 1 | 1765 | 100.0 | 336 | 12 | US-09-995-938A-7 | Sequence 6, Appl1 |
| 2 | 1758 | 99.6 | 335 | 12 | US-09-995-938A-6 | Sequence 7, Appl1 |
| 3 | 1525.5 | 86.5 | 335 | 12 | US-09-995-938A-10 | Sequence 10, Appl1 |
| 4 | 1519.5 | 86.1 | 335 | 12 | US-09-995-938A-8 | Sequence 8, Appl1 |
| 5 | 137.5 | 7.8 | 318 | 12 | US-09-840-746-0 | Sequence 20, Appl1 |
| 6 | 131 | 7.4 | 314 | 15 | US-10-043-487-111 | Sequence 31, Appl1 |
| 7 | 127.5 | 7.2 | 271 | 15 | US-10-105-004-2 | Sequence 2, Appl1 |
| 8 | 127.5 | 7.2 | 271 | 15 | US-10-105-004-121 | Sequence 121, Appl1 |
| 9 | 127.5 | 7.2 | 271 | 15 | US-10-105-004-125 | Sequence 125, Appl1 |
| 10 | 127.5 | 7.2 | 1493 | 9 | US-09-858-754-4 | Sequence 4, Appl1 |
| 11 | 126.5 | 7.2 | 210 | 15 | US-10-106-698-5571 | Sequence 5571, Appl1 |
| 12 | 126.5 | 7.2 | 271 | 15 | US-10-105-004-8 | Sequence 5, Appl1 |
| 13 | 126.5 | 7.2 | 1265 | 15 | US-10-032-585-7168 | Sequence 7168, Appl1 |
| 14 | 124.5 | 7.1 | 271 | 15 | US-10-105-004-131 | Sequence 131, Appl1 |
| 15 | 124 | 7.0 | 1162 | 12 | US-10-234-804-2 | Sequence 2, Appl1 |

| | | | | | | |
|----|-------|-----|-------|----|---------------------|----------------------|
| 16 | 124 | 7.0 | 1.167 | 10 | US-09-801-368-108 | Sequence 108, Appl |
| 17 | 122 | 6.9 | 1.126 | 11 | US-09-917-384-1 | Sequence 1, Appl |
| 18 | 122 | 6.9 | 1.128 | 11 | US-09-917-383-1 | Sequence 1, Appl |
| 19 | 120.5 | 6.8 | 1.193 | 9 | US-09-858-754-3 | Sequence 3, Appl |
| 20 | 120.5 | 6.8 | 1.193 | 14 | US-10-000-864-8 | Sequence 8, Appl |
| 21 | 120 | 6.8 | 1.192 | 9 | US-09-789-386-2 | Sequence 2, Appl |
| 22 | 120 | 6.8 | 1.192 | 9 | US-09-758-140-6 | Sequence 6, Appl |
| 23 | 120 | 6.8 | 1.192 | 9 | US-09-893-348-23 | Sequence 23, Appl |
| 24 | 120 | 6.8 | 1.192 | 9 | US-09-972-859A-2 | Sequence 61, Appl |
| 25 | 120 | 6.8 | 1.192 | 15 | US-10-060-036-71 | Sequence 71, Appl |
| 26 | 119.5 | 6.8 | 1.192 | 15 | US-10-094-466-14 | Sequence 14, Appl |
| 27 | 116.5 | 6.6 | 1.192 | 15 | US-10-155-440-1 | Sequence 1, Appl |
| 28 | 115.5 | 6.5 | 1.192 | 9 | US-09-836-752-6 | Sequence 6, Appl |
| 29 | 114.5 | 6.5 | 1.197 | 12 | US-10-029-386-34403 | Sequence 34203, Appl |
| 30 | 114.5 | 6.5 | 1.197 | 12 | US-09-945-917-57 | Sequence 57, Appl |
| 31 | 114.5 | 6.5 | 1.194 | 12 | US-09-845-917A-57 | Sequence 57, Appl |
| 32 | 114.5 | 6.5 | 1.149 | 11 | US-09-945-917A-58 | Sequence 58, Appl |
| 33 | 114.5 | 6.5 | 1.149 | 12 | US-09-845-917A-58 | Sequence 58, Appl |
| 34 | 114.5 | 6.5 | 1.128 | 11 | US-09-845-917A-3 | Sequence 3, Appl |
| 35 | 114.5 | 6.5 | 1.128 | 12 | US-09-845-917A-3 | Sequence 3, Appl |
| 36 | 114.5 | 6.5 | 1.158 | 11 | US-09-945-917A-4 | Sequence 4, Appl |
| 37 | 114.5 | 6.5 | 1.158 | 12 | US-09-845-917A-4 | Sequence 4, Appl |
| 38 | 113.5 | 6.4 | 1.201 | 15 | US-10-029-386-33339 | Sequence 33339, Appl |
| 39 | 113.5 | 6.4 | 1.201 | 15 | US-10-029-386-33339 | Sequence 33339, Appl |
| 40 | 112 | 6.3 | 1.192 | 12 | US-10-171-311-83 | Sequence 83, Appl |
| 41 | 111.5 | 6.3 | 1.192 | 12 | US-10-163-774-3 | Sequence 3, Appl |
| 42 | 111 | 6.3 | 1.192 | 12 | US-10-163-774-3 | Sequence 3, Appl |
| 43 | 110.5 | 6.3 | 1.192 | 12 | US-10-094-466-16 | Sequence 16, Appl |
| 44 | 110.5 | 6.3 | 1.192 | 12 | US-09-780-511-2 | Sequence 2, Appl |
| 45 | 110.5 | 6.3 | 1.192 | 12 | US-10-409-551-2 | Sequence 2, Appl |
| | | | | | US-10-145-396-12 | Sequence 12, Appl |

ALIGNMENTS

RESULT

| | |
|-----------------------------|-----------------------------------|
| Sequence 7, Application | US/09995938A |
| Publication No. | US20030150026A1 |
| GENERAL INFORMATION: | |
| APPLICANT: | JOANNE CHORY |
| APPLICANT: | ZHIYONG WANG |
| TITLE OF INVENTION: | HGENS INVOLVED IN BRASSINOSTEROID |
| TITLE OF INVENTION: | HORMONE ACTION IN PLANTS |
| FILE REFERENCE: | SALKINS/046A |
| CURRENT APPLICATION NUMBER: | US/09/995,938A |
| CURRENT FILING DATE: | 2001-11-27 |
| NUMBER OF SEQ ID NOS: | 14 |
| SOFTWARE: | FastSeq for Windows Version 4.0 |
| SEQ ID NO 7 | |
| LENGTH: | 336 |
| TYPE: | PRT |
| ORGANISM: | ARABIDOPSIS THALIANA |
| IS-09-995-938A-7 | |

| | | | | |
|-------------|--------|------------|-------|------------|
| Query Match | 100.0% | Score 1765 | DB 12 | Length 336 |
|-------------|--------|------------|-------|------------|

| | | | | | | | | | |
|---------|------|--------------|----|------------|----|--------|----|------|---|
| Matches | 336; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0 |
|---------|------|--------------|----|------------|----|--------|----|------|---|

| QY | 1 | MTSDGATISTAAAAAARRRKPSWRENNRRRENNRRRAVAAKITYTGLRAGDYNLP | 60 |
|----|-----|--|-----|
| Db | 1 | MTSDGATISTAAAAAARRRKPSWRENNRRRENNRRRAVAAKITYTGLRAGDYNLP | 60 |
| QY | 61 | KHCDNNEVLKALCTYBAGVVEEDGTTTRKCGKPLPGEIAGTSRVTPTYSQNOGSLSSAF | 120 |
| Db | 61 | KHCDNNEVLKALCTYBAGVVEEDGTTTRKCGKPLPGEIAGTSRVTPTYSQNOGSLSSAF | 120 |
| QY | 121 | OSPIPSYOVSSSSSFPSRSGEENNNMNSSTPTFFPLRNGTLPSSLPLSLRJSNCCPTVPV | 180 |
| Db | 121 | OSPIPSYOVSSSSSFPSRSGEENNNMNSSTPTFFPLRNGTLPSSLPLSLRJSNCCPTVPV | 180 |
| QY | 181 | SSPTSKNPKPLPNWBSIAKOSMAIAKOSMASPNVPTFAVNSPASPTTRHQHTLATITPBC | 240 |

Db 181 SSPTSNNKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIP 240
Qy 241 DESDSSSTVDSGHWISFOKFAOQOQPFSSASMPPTSPTNLVKPAPQOMSPTTAFOEIGOSS 300
Db 241 DESDSSSTVDSGHWISFOKFAOQOQPFSSASMPPTSPTNLVKPAPQOMSPTTAFOEIGOSS 300
Qy 301 EFKFENSQVKPWEGERIHDVGMEDLELTGNGKARG 336
Db 301 EFKFENSQVKPWEGERIHDVGMEDLELTGNGKARG 336

RESULT 2

US-09-995-938a-6
Sequence 6, Application US/09995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
FILE REFERENCE: SALKINS. 046A
CURRENT APPLICATION NUMBER: US/09/995, 938A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 336
TYPE: PRT
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938a-6

Query Match 99.6%; Score 1758; DB 12; Length 336;
Best Local Similarity 99.7%; Pred. No. 1,2e-133;
Matches 335; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MTSOGATSTSAIAAAAAAAAAARRKPSWRENNRRERRRRRAVAAKIYTGIAOQDYNLP 60
Db 1 MTSOGATSTSAIAAAAAAAAAARRKPSWRENNRRERRRRRAVAAKIYTGIAOQDYNLP 60
Qy 61 KHCDDNEVLKALCEAGVWVEEDGTTTKGCKPLGELAGTSSRVTPYSSONOSPLSSAF 120
Db 61 KHCDDNEVLKALCEAGVWVEEDGTTTKGCKPLGELAGTSSRVTPYSSONOSPLSSAF 120
Qy 121 QSPISYQVSPSSSSPSPSRGEPNNNSSTFPPLRNGGIPSSLSPLRISNSCVTPPV 180
Db 121 QSPISYQVSPSSSSPSPSRGEPNNNSSTFPPLRNGGIPSSLSPLRISNSCVTPPV 180
Qy 181 SSPTSNNKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIP 240
Db 181 SSPTSNNKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIP 240
Qy 241 DESDSSSTVDSGHWISFOKFAOQOQPFSSASMPPTSPTNLVKPAPQOMSPTTAFOEIGOSS 300
Db 241 DESDSSSTVDSGHWISFOKFAOQOQPFSSASMPPTSPTNLVKPAPQOMSPTTAFOEIGOSS 300
Qy 301 EFKFENSQVKPWEGERIHDVGMEDLELTGNGKARG 336
Db 301 EFKFENSQVKPWEGERIHDVGMEDLELTGNGKARG 336

RESULT 3

US-09-995-938a-10
Sequence 10, Application US/09995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
FILE REFERENCE: SALKINS. 046A
CURRENT APPLICATION NUMBER: US/09/995, 938A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 335
TYPE: PRT
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938a-10
Query Match 86.5%; Score 1526.5; DB 12; Length 335;
Best Local Similarity 88.7%; Pred. No. 5.6e-115;
Matches 298; Conservative 10; Mismatches 23; Indels 5; Gaps 4;

Qy 1 MTSOGATSTSAIAAAAAAAAAARRKPSWRENNRRERRRRRAVAAKIYTGIAOQDYNLP 60
Db 1 MTSOGATSTSAIAAAAAAAAAARRKPSWRENNRRERRRRRAVAAKIYTGIAOQDYNLP 59
Qy 61 KHCDDNEVLKALCEAGVWVEEDGTTTKGCKPLGELAGTSSRVTPYSSONOSPLSSAF 120
Db 61 KHCDDNEVLKALCEAGVWVEEDGTTTKGCKPLGELAGTSSRVTPYSSONOSPLSSAF 119
Qy 121 QSPISYQVSPSSSSPSPSRGEPNNNSSTFPPLRNGGIPSSLSPLRISNSCVTPPV 179
Db 120 DSPISYQVSPSSSSPSPSRGEPNNNSSTFPPLRNGGIPSSLSPLRISNSCVTPPV 177
Qy 180 VSSPTSNNKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIP 238
Db 178 VSSPTSNNKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIP 237
Qy 239 ECDESDSSTVDSGHWISFOKFAOQOQPFSSASMPPTSPTNLVKPAPQOMSPTTAFOEIGQ 298
Db 238 ECDESDSSTVDSGHWISFOKFAOQOQPFSSASMPPTSPTNLVKPAPQOMSPTTAFOEIGQ 297
Qy 299 SSEFKFENSQVKPWEGERIHDVGMEDLELTGNGKA 334
Db 298 SSEFKFENSQVKPWEGERIHDVGMEDLELTGNGKA 333

RESULT 4

US-09-995-938a-8
Sequence 8, Application US/09995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
FILE REFERENCE: SALKINS. 046A
CURRENT APPLICATION NUMBER: US/09/995, 938A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 335
TYPE: PRT
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938a-8

Query Match 86.1%; Score 1519.5; DB 12; Length 335;
Best Local Similarity 88.4%; Pred. No. 2.1e-114;
Matches 297; Conservative 10; Mismatches 24; Indels 5; Gaps 4;

Qy 1 MTSOGATSTSAIAAAAAAAAAARRKPSWRENNRRERRRRRAVAAKIYTGIAOQDYNLP 60
Db 1 MTSOGATSTSAIAAAAAAAAAARRKPSWRENNRRERRRRRAVAAKIYTGIAOQDYNLP 59
Qy 61 KHCDDNEVLKALCEAGVWVEEDGTTTKGCKPLGELAGTSSRVTPYSSONOSPLSSAF 120
Db 61 KHCDDNEVLKALCEAGVWVEEDGTTTKGCKPLGELAGTSSRVTPYSSONOSPLSSAF 119
Qy 121 QSPISYQVSPSSSSPSPSRGEPNNNSSTFPPLRNGGIPSSLSPLRISNSCVTPPV 179
Db 120 DSPISYQVSPSSSSPSPSRGEPNNNSSTFPPLRNGGIPSSLSPLRISNSCVTPPV 177
Qy 180 VSSPTSNNKPKPLPWNESIAKOSMAIAKOSMASFNYPFAVASAPSPTRHQFTLATIP 238

| Db | Qy | Db | Qy |
|-----|--|-----|----|
| 178 | VBSPTSRNPKLPTWESFTKOSMSAAAKQSMSTSLNTPFAVAVASAPASPTTHROPHAPATTP | 237 | |
| 239 | ECDSDSSTVDSGHWISFQKFAQQQPFSAWVPTSTFNLVVRPAQOMSPTNAFOEIQ | 298 | |
| 238 | ECDSDSSTVDSGHWISFQKFAQQQPFSAWVPTSTFNLVVRPAQQLSPNTAIOEIQ | 297 | |
| 299 | SSSEKPFNSQVKNPMEGERIHVDGMBDELTTLCNGKA | 334 | |
| 298 | SSSEKPFNSQVKNPMEGERIHVDVAMEDELTTLCNGKA | 333 | |

```

RESULT 5
US-09-840-746-20
Sequence 20, Application US/09840746
Publication No. US20030166501A1
GENERAL INFORMATION:
APPLICANT: Chen, Hui-Mei
APPLICANT: Honchell, Chynthia D.
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Muscin-Related Tumor Marker
FILE REFERENCE: PC-0039 US
CURRENT APPLICATION NUMBER: US/09/840.746
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SBO ID NO. 20
LENGTH: 528
TYPE: PRT
ORGANISM: Sus scrofa
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Genbank ID No. US20030166501A1
US-09-840-746-20

```

| | | | | |
|-----------------------|------------------|-------------------|------------|-------------|
| Query Match | 7.8%; | Score 137.5; | DB 12; | Length 528; |
| Best Local Similarity | 24.6%; | Pred. No. 0.0085; | | |
| Matches 61; | Conservative 32; | Mismatches 88; | Indels 67; | Gaps 9; |

```

99 PLPGG--EIGATSSRVRPYSSSONOPLSSAFOSPIS-----YOVSSPS 133
237 PIPSTTSVQVSSSSSAPTTSATSVQVSSSSSTPIPIPTTSVQVSSSSAPTTSATSVQVSS 296
134 SSFP-----SPKGEPPNNMSSTFPPLRNGIP-SLIPSLRISNCEVTPVSS-- 182
297 SSSPPISSITSVQVSSSSSSPTTSTTSVQVSSSSGSAPTTSATSVQVSSSS--SPISSTI 354
183 ---PTSKNPKPLPNMESIAKOSMAIAKOSMAFNYPYAVSAPASPTHRHQFHLATPIE 239
355 SVQVSSSSSSPTTSTTSVQVSSSSGSAPTTSATSVQVSSSSSVPT-----TSATSVR 405
240 CDESDSSTVDSGHWISQKFAQQQPFASNVPTS-----PTFNLVKKPAQ 284
406 SSSSSSTPIPT-----TTSVQVSSSSSVPTTSATSVQVSSSSSTPIPETTSVQVSSS 457
285 QMSPTTAA 292
458 SSAPTTSA 465

```

RESULT 6
 US-10-043-487-311
 Sequence 311, Application US/10043487
 Publication No. US20030055220A1
 GENERAL INFORMATION:
 APPLICANT: HYBRIGENICS
 APPLICANT: Pierre, LEGRAIN
 TITLE OP INVENTION: Protein-protein interactions between *Shigella flexneri* polypeptides
 TITLE OP INVENTION: mammalian polypeptides
 FILE REFERENCE: B4778A
 CURRENT APPLICATION NUMBER: US/10/043,487
 CURRENT FILING DATE: 2002-04-30
 PRIOR APPLICATION NUMBER: US 60/261,130

```

; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: PatentIn version 3.1.1
; SEQ ID NO 311
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Shigella flexneri
US-10-043-487-311

```

| | | | | |
|--|-------|-----------------|-------|------------|
| Query Match | 7.48 | Score 131 | DB 15 | Length 314 |
| Best Local Similarity | 29.74 | Pred. No. 0.015 | | |
| Matches 65; Conservativs 23; Mismatches 91; Indels 40; Gaps. 12; | | | | |

QY 93 PLGGLGAGTSSRTPTYSQONQSPLESAFOSPSPSYOVSPSSSF--PSPERGEENNNMGSPT 151
Db 108 PTPSPYSPTSPNPTTP--TSPNYSPTSPSYSPPTSPSY--SPTSPSYSSSPSRYPYTP--QSPPT 161
QY 152 FPPFLNLNGGIPSLPLSLRLSNSCPVTPPVYSPTSKNPKP-----LPMWESIA-KOSM 202
Db 162 YTP-----SSPTSPSPS---SPSYSPSPSKYTTPTSPSYSSSSELYPPTSPSKYSTSPKYSBP 213
QY 203 AIAKOSMASFNY-----PRYAVASAP--SPTRHQFHTLATIPCEDESDSSTVD 250
Db 214 TSPKYSPTSPPTSPTPPKYSPTSPSYSPSPYTTPTSPSKYSTPTSPYSPTSPKYSTSPPT 273
QY 251 GHWISFOKFAQOQPPSASNVPTSPFTNLVKAPQOMSPN 289
Db 274 -YSPTSPKSGSTYSPSPSGYSTSPSTPYSLTSPA---TSPD 308

RESULT 7
US-10-105-004-2
; Sequence 2, Application US/10105004
; Publication No. US20030105002A1
; GENERAL INFORMATION:
; Applicant: Microsoft Corporation
; Filing Date: 03/25/2003
; Priority Date: 03/25/2003
; Publication Date: 09/04/2003
; Title: SYSTEM AND METHOD FOR MANAGING
; AND/OR MANIPULATING DATA
; IN A DATABASE
; Abstract: A system and method for managing
; and/or manipulating data in a database
; is provided. The system and method
; includes a database, a user interface,
; and a processor. The user interface
; is configured to receive a request from
; a user to perform a query on the
; database. The processor is configured
; to execute the query on the database
; and to return the results of the query
; to the user interface. The user
; interface is configured to display the
; results of the query to the user.

Query Match 7.2%; Score 127.5; DB 15; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.023;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

QY 6 ATSTSAAAAAAAAAARRKPSRERENRRRRRAVAAKITGRLAAGDYNLPKHCND 65
DB 73 AVMTNLTEARVRVWFKNRRAKRRKERNQOALCKNGFGPO-FNGLMOPYDMPGYSYN 131
QY 66 NEVLKALCVAGWVEEDGTTTRKCKPLPGEIAGTSSRVP-YSSONOSPIS--SAFOS 122
DB 132 NMAAKL-----TSASLSTKSPFFPNSMNVNPLSSQSMFSP 167
QY 123 PIPSYOVSPSSSFPSPSRGEPN-----NMSTFPFLRNGCIPSSLSRISNCPV 176
DB 168 PMSISMSSMSSMVPASVATGVPGSSLNLSNLNLSLSPSL-NSAVPTP-----ACPY 218
QY 177 TPVVSPTSCKNPPPLPWMSIAKOSMAI---AKOSMASFNYPYAVSAPAS 224
DB 219 APPT-----PPVYRDTCNSLSLRLKAKOH--SFGYA--SVOKPAS 258

RESULT 8

US-10-105-004-121
Sequence 121, Application US/10105004
Publication No. US20030105002A1
GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC
AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/105,004
FILING DATE: 22-Mar-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 121:
US-10-105-004-121

Query Match 7.2%; Score 127.5; DB 15; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.023;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

QY 6 ATSTSAAAAAAAAAARRKPSRERENRRRRRAVAAKITGRLAAGDYNLPKHCND 65
DB 73 AVMTNLTEARVRVWFKNRRAKRRKERNQOALCKNGFGPO-FNGLMOPYDMPGYSYN 131

QY 66 NEVLKALCVAGWVEEDGTTTRKCKPLPGEIAGTSSRVP-YSSONOSPIS--SAFOS 122

DB 132 NMAAKL-----TSASLSTKSPFFPNSMNVNPLSSQSMFSP 167

QY 123 PIPSYOVSPSSSFPSPSRGEPN-----NMSTFPFLRNGCIPSSLSRISNCPV 176

DB 168 PMSISMSSMSSMVPASVATGVPGSSLNLSNLNLSLSPSL-NSAVPTP-----ACPY 218

QY 177 TPVVSPTSCKNPPPLPWMSIAKOSMAI---AKOSMASFNYPYAVSAPAS 224

DB 219 APPT-----PPVYRDTCNSLSLRLKAKOH--SFGYA--SVOKPAS 258

RESULT 9

US-10-105-004-125
Sequence 125, Application US/10105004
Publication No. US20030105002A1
GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC
AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/105,004
FILING DATE: 22-Mar-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 125:
US-10-105-004-125

Query Match 7.2%; Score 127.5; DB 15; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.023;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

QY 6 ATSTSAAAAAAAAAARRKPSRERENRRRRRAVAAKITGRLAAGDYNLPKHCND 65
DB 73 AVMTNLTEARVRVWFKNRRAKRRKERNQOALCKNGFGPO-FNGLMOPYDMPGYSYN 131
QY 66 NEVLKALCVAGWVEEDGTTTRKCKPLPGEIAGTSSRVP-YSSONOSPIS--SAFOS 122
DB 132 NMAAKL-----TSASLSTKSPFFPNSMNVNPLSSQSMFSP 167
QY 123 PIPSYOVSPSSSFPSPSRGEPN-----NMSTFPFLRNGCIPSSLSRISNCPV 176
DB 168 PMSISMSSMSSMVPASVATGVPGSSLNLSNLNLSLSPSL-NSAVPTP-----ACPY 218

177 TPVSSPTSKNPKPLNPMESIAKOSMAI-----AKOSMASFNYPFAVASAPAS 224
 219 APPT-----PPVYADTCNSSLASIRLAKQH--SSFGTA--SVQKRPAS 258

US-09-858-754-4
 ; Sequence 4, Application US/09858754
 ; Patent No. US20020055130A1

GENERAL INFORMATION:
 APPLICANT: Johnson, Gary L.
 TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING APOPTOSIS
 FILE REFERENCE: CPI-042
 CURRENT APPLICATION NUMBER: US/09/858,754
 CURRENT FILING DATE: 2001-05-16
 PRIOR APPLICATION NUMBER: 09/023,130
 PRIOR FILING DATE: 1998-02-13
 PRIOR APPLICATION NUMBER: 60/039,740
 PRIOR FILING DATE: 1997-02-14
 NUMBER OF SEQ ID NOS: 13
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 4
 LENGTH: 1493
 TYPE: PRT
 ORGANISM: Rattus norvegicus
 US-09-858-754-4

Query Match 7.2%; Score 127.5; DB 9; Length 1493;
 Best Local Similarity 22.2%; Pred. No. 0.2;
 Matches 76; Conservative 38; Mismatches 124; Indels 99; Gaps 12;

5 GATSTSAAAAAAAAAAARRRKPSRENNRRRRRAVAAKIYTLRAQGDYNLPHKCD 64
 134 GARPPAGAEPPSA-----PSGRMEKN-----ETLKGLHKMD--RPERM 174
 65 NNEVLKALCVEAGVVEEDGTTYRKGCCKPLPEIAGTSSRTVPYSQNSPLSSAFQ 121
 175 IRELTAKTCMPA-WKHEMLERRNRKGPVYVPRIP--IKGDSEMSNLAELQGGQAGSA 231
 122 SPIPSYVSSSSSPSPSRGEPNNMNSSTFPPLRNGIPSLPSLRISNSCPV----- 176
 232 APAKGRSPSPGSSPSGRSGKPS-----PGYRKRVSPVPPQSG 272
 177 --TPVSSPTSKNPKPLNPMESIAKOSMAI-KOSMASFNYPFAVASAPASPTR----- 228
 273 RITPPRAPSPDGPSPYSPETSRVNVKMKARLYLLQOIGPNSFLGQSPDKKIVPI 332
 229 -----HGHTLATITPECESD-----SSTVDS 250
 333 GPONCGCGCTPCILHLFVMLRVQLSPDPMWRKTLKNPEVESLPQKYSRRSSRIKA 392
 251 GHWISFOKFAQO-----QPFASMWPTSPFTNLVYPAPQOMSP 288
 393 PSNRITQKFFVSRMSNCHTTLSSSTSTSSNSISIDEEQONCP 434

US-10-106-698-5571
 ; Sequence 5571, Application US/10106698
 ; Publication No. US20030109690A1

GENERAL INFORMATION:
 APPLICANT: Ruden et al.
 TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
 FILE REFERENCE: PA005PI
 CURRENT APPLICATION NUMBER: US/10/106,698
 CURRENT FILING DATE: 2002-03-27
 PRIOR APPLICATION NUMBER: PCT/US00/26524
 PRIOR FILING DATE: 2000-09-28
 PRIOR APPLICATION NUMBER: US 60/157,137
 PRIOR FILING DATE: 1999-09-29
 PRIOR APPLICATION NUMBER: US 60/163,280
 PRIOR FILING DATE: 1999-11-03

NUMBER OF SEQ ID NOS: 9564
 SOFTWARE: PatentIn Ver. 3.0
 SEQ ID NO 5571
 LENGTH: 210
 TYPE: PRT
 ORGANISM: Homo sapiens

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (137)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: MISC FEATURE
 LOCATION: (171)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: MISC FEATURE
 LOCATION: (179)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: MISC FEATURE
 LOCATION: (193)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: MISC FEATURE
 LOCATION: (197)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: MISC FEATURE
 LOCATION: (210)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-10-106-698-5571 equals any of the naturally occurring L-amino acids

Query Match 7.2%; Score 126.5; DB 15; Length 210;
 Best Local Similarity 25.8%; Pred. No. 0.02;
 Matches 48; Conservative 21; Mismatches 74; Indels 43; Gaps 7;

6 ATSTSAAAAAAAAAAARRRKPSRENNRRRRRAVAAKIYTLRAQGDYNLPHKCDN 65
 38 AWTNLTEARVYVYKNNRAKRRERNOQAEICNGGQPO-FNGLMPYDMDYGVSYN 96
 66 NEVLKALCVEAGVVEEDGTTYRKGCCKPLPEIAGTSSRTVP--YSSQNSPLS--SAFQS 122
 97 NMAAKGL-----TSASLSTKSPFPFNSMWNPLSSQSMFSP 132
 123 PIPEYOVSPSSSPSPSRGEPN-----NNWSTFPPLRNGIPSLPSLRISNSCPV 176
 133 PNSTXSMXSMSSNPPSAVTGVPSSLSLNNLNNLSFXL-NSAVPTX-----ACPY 183
 177 TPVSS 182
 184 APPTSS 189

US-10-105-004-5
 ; Sequence 5, Application US/10105004
 ; Publication No. US20030105002A1

GENERAL INFORMATION:
 APPLICANT: Murray, Jeffrey
 TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC AND DIAGNOSTIC USES THEREOF
 NUMBER OF SEQUENCES: 139
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: POLBY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/105,004
FILING DATE: 22-Mar-2002
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UTA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-105-004-5

Query Match
Best Local Similarity 25.9%; Pred. No. 0.028;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

QY 6 ATSTSAIAAAARRRKPPRRRRRAAAKITYGLRAQDYNLPKHCDN 65
DB 73 AVNTNLTEARVWPKRKRAKRRERNOQALCKNGFGPO-FNGLMQPYDMYGYSTN 131
QY 66 NEVLKALCVAAGWVEEDGTTYRKCKPLPGEIAGTSSRVP-YSSONQSPLS--SAFOS 122
DB 132 NMAAKGL-----TSASLSTKSPFPNPMVNVPLSSQSMFSP 167
QY 123 PISYOVSPSSSPSPSPSGEPN-----NNMSTFPPLRNGGIPSLSLRISNSCPV 176
DB 168 PMSISSMSMSMVPSAVTGVPGSSLSNLTNNLNLSSPSL-NSAVPTP-----ACPY 218
QY 177 TPPVSSPTSKPKPLPNWESIAKQSMI---AKQSMASFVNPFAVASAPAS 224
DB 219 APPT-----PPVYRDTCNSSLASLRUKAKQH--SSPCYA--SVQNPMAS 258

RESULT 13
US-10-032-585-7168
Sequence 7168, Application US/10032585
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone
APPLICANT: Howard, Buseey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
NUMBER OF SEQ ID NOS: 2001-12-20
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7168
LENGTH: 1265
TYPE: PRT
ORGANISM: Candida albicans
US-10-032-585-7168

Query Match
Best Local Similarity 7.2%; Score 126.5; DB 12; Length 1265;
Matches 58; Conservative 34; Mismatches 89; Indels 59; Gaps 9;

QY 93 PLPEIAGTSSRVP-PYS-----SQNOSPSSAPQSPISYOVSPSSSPSPSGEPNN 146
DB 965 PPSISISSTRTTPTSTVPASIKPTPVGNKYAPATSDASQPPSSGPAFTLNSPR 1024
QY 147 NMSSTPPPLRNGIP-----SSLPLRLISNSCPVTPPVSSPTSKPKPLPNWESIAKQSM 202

DB 1025 LAKNPYAPSVTBOPLPKISYATPPAHHLANNGSTPSYAPP--KNPYAVPSTSVSHAGI 1082
QY 203 AIAKQSMASFNPFYVASAPASPTHRHQTTLATIPBCSDSDSTVSGHMISQKRAQO 262
DB 1083 APPPPA-----FKIGSAAPP----- 1099
QY 263 QPFSASM-VPTSPPTNLVKPAPQOM--SPNTAAFOEIGQSEFEKE-NSQVKEPWEGERIH 318
DB 1100 QPFGSSMSMVDPAFNGVPPPPPVGKRAVSTPAAKIEQPPAREPELPVOSKHKGRTH 1159

RESULT 14
US-10-105-004-131
Sequence 131, Application US/10105004
Publication No. US20030105002A1
GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
APPLICANT: Semina, Elena
TITLE OF INVENTION: REG COMPOSITIONS AND THERAPEUTIC
AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & BLIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/105,004
FILING DATE: 22-Mar-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UTA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 131:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 131:
US-10-105-004-131

Query Match
Best Local Similarity 7.1%; Score 124.5; DB 15; Length 271;
Matches 60; Conservative 24; Mismatches 89; Indels 59; Gaps 11;

QY 6 ATSTSAIAAAARRRKPPRRRRRAAAKITYGLRAQDYNLPKHCDN 65
DB 73 AVNTNLTEARVWPKRKRAKRRERNOQALCKNGFGPO-FNGLMQPYDMYGYSTN 131
QY 66 NEVLKALCVAAGWVEEDGTTYRKCKPLPGEIAGTSSRVP-YSSONQSPLS--SAFOS 122
DB 132 NMAAKGL-----TSASLSTKSPFPNPMVNVPLSSQSMFSP 167
QY 123 PISYOVSPSSSPSPSPSGEPN-----NNMSTFPPLRNGGIPSLSLRISNSCPV 176
DB 168 PMSISSMSMSMVPSAVTGVPGSSLSNLTNNLNLSSPSL-NSAVPTP-----ACPY 218
QY 177 TPPVSSPTSKPKPLPNWESIAKQSMI---AKQSMASFVNPFAVASAPAS 224

Db 219 APPT-----PPYVYRDTGNSLSLRLKAKQH-SSFGYA--SVQKPA5 258

RESULT 15

US-10-294-804-2
 / Sequence 2, Application US/10294804
 / Publication No. US20030133948A1
 / GENERAL INFORMATION:
 / APPLICANT: Robertson, Etle S.
 / APPLICANT: Colter, Murray A.
 / TITLE OF INVENTION: Methods to inhibit or enhance the binding of viral DNA
 / TITLE OF INVENTION: to Genomic Host DNA
 / FILE REFERENCE: UM-03778
 / CURRENT APPLICATION NUMBER: US/10/294,804
 / PRIOR FILING DATE: 2002-11-14
 / PRIOR APPLICATION NUMBER: US/09/410,399
 / NUMBER OF SEQ ID NOS: 6
 / SOFTWARE: PatentIn Ver. 2.0
 / SEQ ID NO 2
 / LENGTH: 1162
 / TYPE: PRT
 / ORGANISM: Kaposi's sarcoma-associated herpesvirus
 / US-10-294-804-2

Query Match 7.0%; Score 124; DB 12; Length 1162;
 Best Local Similarity 22.5%; Pred. No. 0.28;

Matches 79; Conservative 49; Mismatches 129; Indels 94; Gaps 17;

QY 8 STSAAAAAAAAARRKPSWRENNRRERRRAVAAKIYGLRAQGDVNLPHKCDNNE 67
 Db 13 STGAPLRGSCGRKRNRPERCGLDGLHLPKRVADSI--DGECCGHTLP----- 63
 QY 68 VLKALVEAGWVVEEDGTYRKCKPLGEIAGTSRVTPYSSQNS-----PLSS 118
 Db 64 -----IPGSPVFTSG--LPAFVSPPLVAPISPAPATPLPPALLPPT 108
 QY 119 AFQSPF-PSYQVSSSSSFPSPSRG-EPNNMSSTFPFLNNGIPLSLPRLISNCPV 176
 Db 109 TSSSPFIFSHVSPGTDTHSPSPALPPTQSPSSQRP-----PLSSPGRPDSSTPM 161
 QY 177 TPPVS-----SPTSKNPKP--LPNWSIAKSM-AIAKSMASFNYPF---YAVSAP 222
 Db 162 RPPPSQOTTPPHSTTPPPEPSPKSSPDSLAPSTLRLKRRLLSPQGPSTLNPICQSP 221
 QY 223 ASPTHRHQFTLATIPKCDSDSSTVDGHWISFOKPAQOQPF---SAS-----MVFT 272
 Db 222 VSD-----PRCDFANRNV-----YPPWATESPIYVSSSDGDTTPRPQPT 261
 QY 273 SP-TFNLVKRPAPOOMSPNTAFA-----OEIGSSSEKFNENSVKPKWEGE 315
 Db 262 SPISISSSSPSGSGWGDITMLVLLAEIAEASAKNEKESENNQAGEDNGD 312

Search completed: December 19, 2003, 14:32:05
 Job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2003, 14:24:08 / Search time 43 Seconds

(without alignments)
1240.282 Million cell updates/sec

Title: US-09-995-938a-7

Perfect score: 1765
Sequence: 1 MTSQATSTSTAAAAA.....IHDVMEDELTLGKXRG 336

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 aeqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB aeq length: 0
Maximum DB aeq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: A_GeneSeq_19JUN03.*

1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------|
| 1 | 1758 | 99.6 | 336 | 21 | AAAG16755 |
| 2 | 1511.5 | 85.6 | 335 | 21 | AAAG14237 |
| 3 | 1511.5 | 85.6 | 335 | 21 | AAAG15884 |
| 4 | 1447 | 82.0 | 318 | 21 | AAAG14238 |
| 5 | 1447 | 82.0 | 318 | 21 | AAAG15885 |
| 6 | 1052 | 59.6 | 239 | 21 | AAAG14239 |
| 7 | 1052 | 59.6 | 239 | 21 | AAAG15886 |
| 8 | 994 | 56.3 | 189 | 21 | AAAG16756 |
| 9 | 541 | 30.7 | 284 | 21 | AAAG30139 |

| | | | | | | |
|----|-------|------|------|----|-----------|----------------------|
| 10 | 541 | 30.7 | 284 | 21 | AAAG27278 | Arabidopsis thaliana |
| 11 | 330 | 18.7 | 169 | 21 | AAAG29487 | Arabidopsis thaliana |
| 12 | 235.5 | 13.3 | 205 | 21 | AAAG30141 | Arabidopsis thaliana |
| 13 | 235.5 | 13.3 | 205 | 21 | AAAG2730 | Arabidopsis thaliana |
| 14 | 235.5 | 13.3 | 208 | 21 | AAAG30140 | Arabidopsis thaliana |
| 15 | 235.5 | 13.3 | 208 | 21 | AAAG2729 | Arabidopsis thaliana |
| 16 | 203.5 | 11.5 | 687 | 23 | ABBS92131 | Arabidopsis thaliana |
| 17 | 190 | 10.8 | 689 | 23 | ABBS92130 | Arabidopsis thaliana |
| 18 | 131 | 7.4 | 314 | 23 | ABSG70137 | Arabidopsis thaliana |
| 19 | 127 | 7.2 | 297 | 22 | AAAG74609 | Human prey protein |
| 20 | 126.5 | 7.2 | 210 | 22 | AAAG74797 | Human prey protein |
| 21 | 126.5 | 7.2 | 1265 | 23 | ABP73331 | Human prey protein |
| 22 | 126 | 7.1 | 904 | 22 | ABSG71477 | Human prey protein |
| 23 | 126 | 7.1 | 904 | 22 | ABSG70014 | Human prey protein |
| 24 | 124.5 | 7.1 | 1887 | 22 | ABBS82245 | Human prey protein |
| 25 | 124 | 7.0 | 1162 | 21 | AAV96255 | Human prey protein |
| 26 | 124 | 7.0 | 1162 | 21 | AAV96255 | Human prey protein |
| 27 | 124 | 7.0 | 1162 | 22 | AAV96255 | Human prey protein |
| 28 | 124 | 7.0 | 1162 | 22 | AAV96255 | Human prey protein |
| 29 | 123.5 | 7.0 | 259 | 21 | AAAG32033 | Human prey protein |
| 30 | 123.5 | 7.0 | 268 | 21 | AAAG32032 | Human prey protein |
| 31 | 123 | 7.0 | 989 | 21 | AAAG37223 | Human prey protein |
| 32 | 123 | 7.0 | 994 | 21 | AAAG37222 | Human prey protein |
| 33 | 123 | 7.0 | 1006 | 21 | AAAG37221 | Human prey protein |
| 34 | 122.5 | 6.9 | 1549 | 22 | ABBS5583 | Human prey protein |
| 35 | 122.5 | 6.9 | 2414 | 16 | AAAG4882 | Human prey protein |
| 36 | 122.5 | 6.9 | 2414 | 19 | AAAG40057 | Human prey protein |
| 37 | 122.5 | 6.9 | 2414 | 19 | AAAG40057 | Human prey protein |
| 38 | 122.5 | 6.9 | 2414 | 23 | ABBS06340 | Human prey protein |
| 39 | 122.5 | 6.9 | 2414 | 24 | ABBS03968 | Human prey protein |
| 40 | 122.5 | 6.9 | 2414 | 24 | ABBS03975 | Human prey protein |
| 41 | 122.5 | 6.9 | 2414 | 24 | ABBS03976 | Human prey protein |
| 42 | 122.5 | 6.9 | 2414 | 24 | ABBS03978 | Human prey protein |
| 43 | 122 | 6.9 | 2414 | 24 | ABBS03980 | Human prey protein |
| 44 | 122 | 6.9 | 1475 | 22 | ABBS73029 | Human prey protein |
| 45 | 122 | 6.9 | 1475 | 22 | ABBS71451 | Human prey protein |
| | | | 3238 | 22 | ABBS71715 | Human prey protein |

ALIGNMENTS

| | | |
|----------|--|------------------------------------|
| RESULT 1 | AAAG16755 | Standard; Protein; 336 AA. |
| ID | AAAG16755 | |
| AC | AAAG16755 | |
| DT | 17-OCT-2000 (first entry) | |
| XX | Arabidopsis thaliana | protein fragment SEQ ID NO: 17522. |
| DE | Arabidopsis thaliana | protein fragment SEQ ID NO: 17522. |
| XX | Arabidopsis thaliana | protein fragment SEQ ID NO: 17522. |
| KW | Protein identification; signal transduction pathway; metabolic pathway; hybridization assay; genetic mapping; gene expression control; promoter; termination sequence. | |
| XX | Arabidopsis thaliana | |
| OS | Arabidopsis thaliana | |
| XX | Arabidopsis thaliana | |
| PN | EP1033405-A2. | |
| XX | EP1033405-A2. | |
| PD | 06-SEP-2000. | |
| XX | 06-SEP-2000. | |
| PF | 25-FEB-2000; 2000EP-0301439. | |
| XX | 25-FEB-2000; 2000EP-0301439. | |
| XX | 25-FEB-1999; 99US-0121825. | |
| PR | 05-MAR-1999; 99US-0121825. | |
| PR | 09-MAR-1999; 99US-0123548. | |
| PR | 23-MAR-1999; 99US-0125788. | |
| PR | 25-MAR-1999; 99US-0126264. | |
| PR | 29-MAR-1999; 99US-0126785. | |
| PR | 01-APR-1999; 99US-0127462. | |
| PR | 06-APR-1999; 99US-0128234. | |
| PR | 08-APR-1999; 99US-0128714. | |
| PR | 16-APR-1999; 99US-0129845. | |

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0133048.
PR 30-APR-1999; 99US-0133407.
PR 04-MAY-1999; 99US-0134884.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140883.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142330.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143622.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158359.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 99.6%; Score 1758; DB 21; Length 336;
Best Local Similarity 99.7%; Pred. No. 1,56-142;
Matches 335; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MTSDBGTSTBAAAAAAAAAARRRPSWRENNRRRRRAVAKIYTGIRAGDYNLP 60
DB 1 MTSDBGTSTBAAAAAAAAAARRRPSWRENNRRRRRAVAKIYTGIRAGDYNLP 60
OY 61 KICDNEVVKALCEVAGWVEEDGTYRKCKPLPGIAGTSRVTTPYSSQNSPLSSAF 120
DB 61 KICDNEVVKALCEVAGWVEEDGTYRKCKPLPGIAGTSRVTTPYSSQNSPLSSAF 120
OY 121 QSPISPYOVSPSSSP 180
DB 121 QSPISPYOVSPSSSP 180
OY 121 QSPISPYOVSPSSSP 180
DB 121 QSPISPYOVSPSSSP 180
OY 181 SEPTSNPKPLPNWESIAKOSMAIAKOSMAIAKOSMAIAKOSMAIAKOSMAIAKOSMA 240
DB 181 SEPTSNPKPLPNWESIAKOSMAIAKOSMAIAKOSMAIAKOSMAIAKOSMAIAKOSMA 240
OY 241 DESDSTVDSGHWISFOKFAQQOFPSSASVPTSPFTNLVPAAPQMSPTAFAQZIGQS 300
DB 241 DESDSTVDSGHWISFOKFAQQOFPSSASVPTSPFTNLVPAAPQMSPTAFAQZIGQS 300
OY 301 EFKFENSQVKPWEGERIHVDGMEDELTLTGNGKARG 336
DB 301 EFKFENSQVKPWEGERIHVDGMEDELTLTGNGKARG 336

RESULT 2

AAAG14237 standard; Protein; 335 AA.

AAAG14237;

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 14022.

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129847.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-014342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147360.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.

PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.

Query Match 85.6%; Score 1511.5; DB 21; Length 335;
Best Local Similarity 88.1%; Pred. No. 2, 2e-121;
Matches 296; Conservative 10; Mismatches 25; Indels 5; Gaps 4;

QY 1 MTSDDATSTSAALAAAAAARRRRRRRRRRRRAVAAYAKIYGLRAGDYNLP 60
DB 1 MTSDDATSTSAALAAAAAATRRKPSWRERENRRRRRRRAVAAYAKIYGLRAGDYNLP 59
QY KCDNNEVLKALCSEAGWVEEDGTYRKCKPLPEIAGTSSRVTPYSSONQSPLSAF 120
DB KCDNNEVLKALCSEAGWVEEDGTYRKCKPLPEIAGTSSRVTPYSSONQSPLSAF 119
QY 121 OSPIPSYOVSPSSSFPSPSR-GEPNNNMSTFPPLRNCGIFSSLSPLAISNCPVTPP 179
DB 120 DSPILSYOVSPSSSFPSPSRVDPNNH--ISTIFFPLRNCGIFSSLSPLAISNCPVTPP 177
QY 180 VSSPTSKNRPPLPWNMSIAKOSMAI-AKOSMAFPNPFYVSA-PASPTHRHOPFLTIIP 238
DB 178 VSSPTSKNRPPLPWNMSIAKOSMAI-AKOSMAFPNPFYVSA-PASPTHRHOPFLTIIP 237
QY 239 ECDSDSSTYDSGHWISFOKFAOQOPFSASMTVPTFTFNLVKPAPOQMSPTNTAFOEIQ 298
DB 238 ECDSDSSTYDSGHWISFOKFAOQOPFSASMTVPTFTFNLVKPAPOQMSPTNTAFOEIQ 297
QY 299 SSEPKFENGQVKNWEGERRIHVQMEDLETLGNKA 334
DB 298 SSEPKFENGQVKNWEGERRIHVQMEDLETLGNKA 333

RESULT 3
AAG15884
ID AAG15884 standard; Protein; 335 AA.
XX
AC AAG15884;
XX

DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 16311.
DE Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
PN EPI033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX PF
XX 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0133180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132464.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 06-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134841.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135523.
XX PR 24-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.
XX PR 27-MAY-1999; 99US-0136392.
XX PR 28-MAY-1999; 99US-0136782.
XX PR 01-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 04-JUN-1999; 99US-0137724.
XX PR 07-JUN-1999; 99US-0137824.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 14-JUN-1999; 99US-0139119.
XX PR 16-JUN-1999; 99US-0139452.
XX PR 17-JUN-1999; 99US-0139492.
XX PR 18-JUN-1999; 99US-0139454.
XX PR 18-JUN-1999; 99US-0139456.
XX PR 18-JUN-1999; 99US-0139457.
XX PR 18-JUN-1999; 99US-0139458.
XX PR 18-JUN-1999; 99US-0139459.
XX PR 18-JUN-1999; 99US-0139460.
XX PR 18-JUN-1999; 99US-0139461.
XX PR 18-JUN-1999; 99US-0139462.
XX PR 18-JUN-1999; 99US-0139463.
XX PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143544.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145145.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 25-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.


```

PR 27-AUG-1999; 99US-0151066
PR 27-AUG-1999; 99US-0151080
PR 30-AUG-1999; 99US-0151303
PR 31-AUG-1999; 99US-0151308
PR 01-SEP-1999; 99US-0151330
PR 07-SEP-1999; 99US-0152363
PR 10-SEP-1999; 99US-0153070
PR 13-SEP-1999; 99US-0153758
PR 15-SEP-1999; 99US-0154018
PR 16-SEP-1999; 99US-0154039
PR 20-SEP-1999; 99US-0154779
PR 22-SEP-1999; 99US-0155139
PR 23-SEP-1999; 99US-0155486
PR 24-SEP-1999; 99US-0155659
PR 28-SEP-1999; 99US-0156458
PR 29-SEP-1999; 99US-0156596
PR 04-OCT-1999; 99US-0157117
PR 05-OCT-1999; 99US-0157753
PR 06-OCT-1999; 99US-0157865
PR 07-OCT-1999; 99US-0158029
PR 08-OCT-1999; 99US-0158232
PR 12-OCT-1999; 99US-0158369
PR 13-OCT-1999; 99US-0159293
PR 13-OCT-1999; 99US-0159294
PR 14-OCT-1999; 99US-0159329
PR 14-OCT-1999; 99US-0159330
PR 14-OCT-1999; 99US-0159331
PR 14-OCT-1999; 99US-0159637
PR 14-OCT-1999; 99US-0159638
PR 18-OCT-1999; 99US-0159584
PR 21-OCT-1999; 99US-0160741
PR 21-OCT-1999; 99US-0160767
PR 21-OCT-1999; 99US-0160768
PR 21-OCT-1999; 99US-0160770
PR 21-OCT-1999; 99US-0160814
PR 21-OCT-1999; 99US-0160815
PR 22-OCT-1999; 99US-0160980
PR 22-OCT-1999; 99US-0160981
PR 22-OCT-1999; 99US-0160989
PR 22-OCT-1999; 99US-0161404
PR 22-OCT-1999; 99US-0161405
PR 22-OCT-1999; 99US-0161406
PR 26-OCT-1999; 99US-0161359
PR 26-OCT-1999; 99US-0161360
PR 26-OCT-1999; 99US-0161361
PR 28-OCT-1999; 99US-0161920
PR 28-OCT-1999; 99US-0161922
PR 28-OCT-1999; 99US-0161992
PR 29-OCT-1999; 99US-0162142
PR 29-OCT-1999; 99US-0162147

```

Query Match 85.6%; Score 1511.5; DB 21; Length 335;
 Best Local Similarity 88.1%; Pred. No. 2.2e-121; Indels 5; Gaps 4;
 Matches 296; Conservative 10; Mismatches 25;

```

Qy 1 MTS D G A T S T S A A A A A A A A A A R K P S W R E R E N R R R R R A V A A K I Y T G L A Q D Y N L P 60
Db 1 M T S G A I S T S - A A A A A A A A M A T R R K S W R E R E N R R R R R A V A A K I Y T G L A Q D Y N L P 59
Qy 61 K H C D N E V L K A L C V E A G V N V E D G T T Y R K G C P L P G E I A G T S S R T P Y S S Q N S Q S L S N A F 120
Db 60 K H C D N E V L K A L C S E A G V N V E D G T T Y R K G H K P L P G D M A G S S S R A T P Y S S H N Q S P L S S T F 119
Qy 121 Q S P I P S Y V S P S S S F P S P S R - G E P N N M S S T F P P L R N G G I P S L P L R I S N S C P V T P P 179
Db 120 D S P I L S Y V S P S S S F P S P S R G D P H N - I S T I P P L R N G G I P S L P L R I S N S A P V T P P 177
Qy 180 V S S P T S K N P K P L P N N E S I A K S M A I - A K S M A S F N Y P F Y A V A S A P A S P T H R Q F H T A T I P 238
Db 178 V S S P T S R N S K P L P T E S I P T K O S M A A K S M T S L V Y P F A V A S A P A S P T H R Q F H A P A T I P 237
Qy 239 E C D E S D S T V D S G H W I S F O K F A Q Q O P F S A S M V P T S P T F L V K P A Q O S P N T A A O E I Q 298

```

```

Db 238 E C D E S D S T V D S G H W I S F O K F A Q Q O P F S A S M V P T S P T F L V K P A Q O L S P N T A A O E I Q 297
Qy 299 S S E F K F E N S O V K P W E G E R I H D V G M E D L E L T L G N G K A 334
Db 298 S S E F K F E N S O V K P W E G E R I H D V A M E D L E L T L G N G K A 333

RESULT 4
AAG14238
ID AAG14238 standard; Protein; 318 AA.
XX
AC AAG14238;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 14023.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence
XX
OS Arabidopsis thaliana
XX
PN EP1033405-A2
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0901439.
XX
PR 25-FEB-1999; 99US-0121825
PR 05-MAR-1999; 99US-0123180
PR 09-MAR-1999; 99US-0123548
PR 23-MAR-1999; 99US-0125788
PR 25-MAR-1999; 99US-0126284
PR 29-MAR-1999; 99US-0126785
PR 01-APR-1999; 99US-0127462
PR 06-APR-1999; 99US-0128234
PR 08-APR-1999; 99US-0128714
PR 16-APR-1999; 99US-0129845
PR 19-APR-1999; 99US-0130077
PR 21-APR-1999; 99US-0130449
PR 23-APR-1999; 99US-0130510
PR 23-APR-1999; 99US-0130891
PR 28-APR-1999; 99US-0131449
PR 30-APR-1999; 99US-0132048
PR 30-APR-1999; 99US-0132407
PR 04-MAY-1999; 99US-0132484
PR 05-MAY-1999; 99US-0132485
PR 06-MAY-1999; 99US-0132486
PR 06-MAY-1999; 99US-0132487
PR 07-MAY-1999; 99US-0132863
PR 11-MAY-1999; 99US-0134256
PR 14-MAY-1999; 99US-0134218
PR 14-MAY-1999; 99US-0134219
PR 14-MAY-1999; 99US-0134221
PR 14-MAY-1999; 99US-0134370
PR 14-MAY-1999; 99US-0134376
PR 18-MAY-1999; 99US-0134941
PR 19-MAY-1999; 99US-0135124
PR 20-MAY-1999; 99US-0135353
PR 21-MAY-1999; 99US-0135629
PR 24-MAY-1999; 99US-0136021
PR 25-MAY-1999; 99US-0136392
PR 27-MAY-1999; 99US-0136782
PR 28-MAY-1999; 99US-0137222
PR 01-JUN-1999; 99US-0137528
PR 03-JUN-1999; 99US-0137502
PR 04-JUN-1999; 99US-0137724
PR 07-JUN-1999; 99US-0138094
PR 08-JUN-1999; 99US-0138540
PR 10-JUN-1999; 99US-0138847
PR 10-JUN-1999; 99US-0139113
PR 16-JUN-1999; 99US-0139452

```


DB 62 VERDGTTRKGRHPLQDMKSSSRATPVSSHNQSPISFDPDPITISYQVSPSSSPSP 121
QY 140 SR-GEPPNNMSTFPPLRNGCIPSSLPRLISNSCPVPVSPTRSKPKPLNNESIA 198
DB 122 SRGDDHAN--ISITFPLRNGCIPSSLPRLISNSAVTTPVSPTRSKPKPLPTWESFT 179
QY 199 KQSMAL-AKQSMASFNVPFYAVASAPASPTRHQFHTLATIPECDSSDSTVDGHWISFQ 257
DB 180 KQSMMAAKQSMSTSLNVPFYAVASAPASPTHROFHAFATIPEDCDSSDSTVDGHWISFQ 239
QY 258 KFAQOQPFSSAMVPTSPTEFLVVKPAPQOMSPTTAFOIGOSSSEKPKENSQVPMEGERI 317
DB 240 KFAQOQPFSSAMVPTSPTEFLVVKPAPQOLSPTTAFOIGOSSSEKPKENSQVPMEGERI 299
QY 318 HDVGMEDLETLNGKA 334
DB 300 HDVAMEDLETLNGKA 316

RESULT 5
AAG15885
ID AAG15885 standard; Protein: 318 AA..
XX AAG15885;
AC
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 16312.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; .99US-0121825.
PR 05-MAR-1999; .99US-0123180.
PR 09-MAR-1999; .99US-0123548.
PR 23-MAR-1999; .99US-0125788.
PR 23-MAR-1999; .99US-0126264.
PR 29-MAR-1999; .99US-0126785.
PR 01-APR-1999; .99US-0127462.
PR 06-APR-1999; .99US-0128234.
PR 08-APR-1999; .99US-0128714.
PR 16-APR-1999; .99US-0129845.
PR 19-APR-1999; .99US-0130077.
PR 21-APR-1999; .99US-0130449.
PR 23-APR-1999; .99US-0130510.
PR 23-APR-1999; .99US-0130891.
PR 28-APR-1999; .99US-0131449.
PR 30-APR-1999; .99US-0132048.
PR 30-APR-1999; .99US-0132407.
PR 04-MAY-1999; .99US-0132484.
PR 05-MAY-1999; .99US-0132485.
PR 06-MAY-1999; .99US-0132486.
PR 06-MAY-1999; .99US-0132487.
PR 07-MAY-1999; .99US-0132863.
PR 11-MAY-1999; .99US-0134256.
PR 14-MAY-1999; .99US-0134218.
PR 14-MAY-1999; .99US-0134219.
PR 14-MAY-1999; .99US-0134221.
PR 14-MAY-1999; .99US-0134370.
PR 18-MAY-1999; .99US-0134768.
PR 19-MAY-1999; .99US-0134941.
PR 20-MAY-1999; .99US-0135124.
PR 21-MAY-1999; .99US-0135353.

PR 24-MAY-1999; .99US-0135629.
PR 25-MAY-1999; .99US-0136021.
PR 27-MAY-1999; .99US-0136392.
PR 28-MAY-1999; .99US-0136782.
PR 01-JUN-1999; .99US-0137222.
PR 03-JUN-1999; .99US-0137528.
PR 04-JUN-1999; .99US-0137502.
PR 07-JUN-1999; .99US-0137724.
PR 08-JUN-1999; .99US-0138094.
PR 10-JUN-1999; .99US-0138540.
PR 10-JUN-1999; .99US-0138847.
PR 14-JUN-1999; .99US-0139119.
PR 16-JUN-1999; .99US-0139452.
PR 16-JUN-1999; .99US-0139453.
PR 17-JUN-1999; .99US-0139492.
PR 18-JUN-1999; .99US-0139454.
PR 18-JUN-1999; .99US-0139455.
PR 18-JUN-1999; .99US-0139456.
PR 18-JUN-1999; .99US-0139457.
PR 18-JUN-1999; .99US-0139458.
PR 18-JUN-1999; .99US-0139459.
PR 18-JUN-1999; .99US-0139460.
PR 18-JUN-1999; .99US-0139461.
PR 18-JUN-1999; .99US-0139462.
PR 18-JUN-1999; .99US-0139463.
PR 18-JUN-1999; .99US-0139750.
PR 18-JUN-1999; .99US-0139753.
PR 18-JUN-1999; .99US-0139817.
PR 21-JUN-1999; .99US-0139899.
PR 22-JUN-1999; .99US-0140353.
PR 23-JUN-1999; .99US-0140354.
PR 23-JUN-1999; .99US-0140695.
PR 24-JUN-1999; .99US-0140823.
PR 28-JUN-1999; .99US-0140921.
PR 29-JUN-1999; .99US-0141287.
PR 30-JUN-1999; .99US-0141842.
PR 01-JUL-1999; .99US-0142154.
PR 02-JUL-1999; .99US-0142305.
PR 06-JUL-1999; .99US-0142803.
PR 08-JUL-1999; .99US-0142920.
PR 09-JUL-1999; .99US-0142977.
PR 12-JUL-1999; .99US-0143542.
PR 14-JUL-1999; .99US-0143624.
PR 15-JUL-1999; .99US-0144005.
PR 16-JUL-1999; .99US-0144085.
PR 16-JUL-1999; .99US-0144325.
PR 19-JUL-1999; .99US-0144331.
PR 19-JUL-1999; .99US-0144332.
PR 19-JUL-1999; .99US-0144333.
PR 19-JUL-1999; .99US-0144334.
PR 19-JUL-1999; .99US-0144335.
PR 20-JUL-1999; .99US-0144352.
PR 20-JUL-1999; .99US-0144632.
PR 20-JUL-1999; .99US-0144884.
PR 20-JUL-1999; .99US-0144814.
PR 21-JUL-1999; .99US-0145086.
PR 21-JUL-1999; .99US-0145088.
PR 22-JUL-1999; .99US-0145087.
PR 22-JUL-1999; .99US-0145089.
PR 22-JUL-1999; .99US-0145192.
PR 22-JUL-1999; .99US-0145192.
PR 23-JUL-1999; .99US-0145218.
PR 23-JUL-1999; .99US-0145224.
PR 26-JUL-1999; .99US-0145276.
PR 27-JUL-1999; .99US-0145913.
PR 27-JUL-1999; .99US-0145918.
PR 27-JUL-1999; .99US-0145919.
PR 28-JUL-1999; .99US-0145951.
PR 02-AUG-1999; .99US-0146386.
PR 02-AUG-1999; .99US-0146388.

```

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147426.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157717.
PR 05-OCT-1999; 99US-0157865.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

```

```

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Beet Local Similarity 82.0%; Score 1447; DB 21; Length 318;
Matches 279; Conservative 10; Mismatches 24; Indels 4; Gaps 3;

QY 20 AARRKPSWREERNRRRRRAVAKIYTGIRAGQADNLPBGCDNNEVLKALCVAGNV 79
DB 2 ATRKPSWREERNRRRRRAVAKIYTGIRAGQADNLPBGCDNNEVLKALCVAGNV 61
QY 80 VEEDGTYRKCKPLPGEIAGTSSRVTPYSSONOSPLSAFOSPIPSYOVSSSSFFSP 139
DB 62 VEEDGTYRKCKPLPGEIAGTSSRVTPYSSONOSPLSAFOSPIPSYOVSSSSFFSP 121
QY 140 SR-GEPNMMSSTFPPLRNGCIPSSLPRLISNSCPVTPYSSPTSKPKPLPWESIA 198
DB 122 SRVDPHN--ISTFPLRNGCIPSSLPRLISNSHPVTPYSSPTSKPKPLPWESFT 179
QY 199 KOSMAI-AKOSMASFNYPFAVASAPASPTHRHOFHTLATIPECDSDSSTVDGHWISFO 257
DB 180 KOSMASAKOSMISLNTFYAVASAPASPTHRHOFHTLATIPECDSDSSTVDGHWISFO 239
QY 258 KFAQQCPFSASMYPTSPFNLVKPAPOOMSPTAFAOBIIGOSSEPKFENSQVKPMEGERI 317
DB 240 KFAQQCPFSASMYPTSPFNLVKPAPOOLSPTAAJDEIGOSSEPKFENSQVKPMEGERI 299
QY 318 HDVGMEDLETLGNGKA 334
DB 300 HDVAMEDLETLGNGKA 316

RESULT 6
AAAG14239
ID AAAG14239 standard; Protein: 239 AA.
XX AC
XX AAAG14239;
XX DT
XX 17-OCT-2000 (first entry)
XX DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 14024.
XX KW
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW
XX Arabidopsis thaliana.
XX OS
XX Arabidopsis thaliana.
XX PN
XX EP1033405-A2.
XX PD
XX 06-SEP-2000.
XX PF
XX 25-FEB-2000; 2000EP-0301439.
XX PR
XX 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0133484.

```

| | | | | | |
|----|--------------|---------------|----|--------------|---------------|
| PR | 05-MAY-1999; | 99US-0132485; | PR | 22-JUL-1999; | 99US-0145087; |
| PR | 06-MAY-1999; | 99US-0132486; | PR | 23-JUL-1999; | 99US-0145089; |
| PR | 06-MAY-1999; | 99US-0132487; | PR | 23-JUL-1999; | 99US-0145192; |
| PR | 07-MAY-1999; | 99US-0132863; | PR | 23-JUL-1999; | 99US-0145145; |
| PR | 11-MAY-1999; | 99US-0134256; | PR | 23-JUL-1999; | 99US-0145218; |
| PR | 14-MAY-1999; | 99US-0134218; | PR | 23-JUL-1999; | 99US-0145224; |
| PR | 14-MAY-1999; | 99US-0134219; | PR | 26-JUL-1999; | 99US-0145276; |
| PR | 14-MAY-1999; | 99US-0134221; | PR | 27-JUL-1999; | 99US-0145913; |
| PR | 14-MAY-1999; | 99US-0134370; | PR | 27-JUL-1999; | 99US-0145918; |
| PR | 18-MAY-1999; | 99US-0134768; | PR | 27-JUL-1999; | 99US-0145991; |
| PR | 19-MAY-1999; | 99US-0134941; | PR | 28-JUL-1999; | 99US-0146386; |
| PR | 20-MAY-1999; | 99US-0135124; | PR | 02-AUG-1999; | 99US-0146388; |
| PR | 21-MAY-1999; | 99US-0135353; | PR | 02-AUG-1999; | 99US-0146389; |
| PR | 24-MAY-1999; | 99US-0135629; | PR | 03-AUG-1999; | 99US-0147038; |
| PR | 25-MAY-1999; | 99US-0136021; | PR | 04-AUG-1999; | 99US-0147204; |
| PR | 27-MAY-1999; | 99US-0136392; | PR | 04-AUG-1999; | 99US-0147302; |
| PR | 28-MAY-1999; | 99US-0136782; | PR | 05-AUG-1999; | 99US-0147312; |
| PR | 01-JUN-1999; | 99US-0137222; | PR | 05-AUG-1999; | 99US-0147260; |
| PR | 03-JUN-1999; | 99US-0137528; | PR | 06-AUG-1999; | 99US-0147303; |
| PR | 04-JUN-1999; | 99US-0137502; | PR | 06-AUG-1999; | 99US-0147416; |
| PR | 07-JUN-1999; | 99US-0137724; | PR | 09-AUG-1999; | 99US-0147493; |
| PR | 08-JUN-1999; | 99US-0138094; | PR | 09-AUG-1999; | 99US-0147935; |
| PR | 10-JUN-1999; | 99US-0138540; | PR | 10-AUG-1999; | 99US-0148111; |
| PR | 10-JUN-1999; | 99US-0138847; | PR | 11-AUG-1999; | 99US-0148319; |
| PR | 14-JUN-1999; | 99US-0139119; | PR | 12-AUG-1999; | 99US-0148341; |
| PR | 16-JUN-1999; | 99US-0139452; | PR | 13-AUG-1999; | 99US-0148565; |
| PR | 16-JUN-1999; | 99US-0139453; | PR | 13-AUG-1999; | 99US-0148664; |
| PR | 17-JUN-1999; | 99US-0139492; | PR | 16-AUG-1999; | 99US-0149368; |
| PR | 18-JUN-1999; | 99US-0139454; | PR | 17-AUG-1999; | 99US-0149115; |
| PR | 18-JUN-1999; | 99US-0139455; | PR | 18-AUG-1999; | 99US-0149426; |
| PR | 18-JUN-1999; | 99US-0139456; | PR | 20-AUG-1999; | 99US-0149723; |
| PR | 18-JUN-1999; | 99US-0139457; | PR | 20-AUG-1999; | 99US-0149723; |
| PR | 18-JUN-1999; | 99US-0139458; | PR | 20-AUG-1999; | 99US-0149929; |
| PR | 18-JUN-1999; | 99US-0139459; | PR | 23-AUG-1999; | 99US-0149902; |
| PR | 18-JUN-1999; | 99US-0139460; | PR | 23-AUG-1999; | 99US-0149930; |
| PR | 18-JUN-1999; | 99US-0139461; | PR | 25-AUG-1999; | 99US-0150566; |
| PR | 18-JUN-1999; | 99US-0139462; | PR | 26-AUG-1999; | 99US-0150884; |
| PR | 18-JUN-1999; | 99US-0139463; | PR | 27-AUG-1999; | 99US-0151065; |
| PR | 18-JUN-1999; | 99US-0139750; | PR | 27-AUG-1999; | 99US-0151066; |
| PR | 18-JUN-1999; | 99US-0139763; | PR | 27-AUG-1999; | 99US-0151080; |
| PR | 21-JUN-1999; | 99US-0139817; | PR | 30-AUG-1999; | 99US-0151303; |
| PR | 22-JUN-1999; | 99US-0139899; | PR | 31-AUG-1999; | 99US-0151438; |
| PR | 23-JUN-1999; | 99US-0140053; | PR | 01-SEP-1999; | 99US-0151930; |
| PR | 23-JUN-1999; | 99US-0140054; | PR | 07-SEP-1999; | 99US-0152363; |
| PR | 24-JUN-1999; | 99US-0140695; | PR | 10-SEP-1999; | 99US-0153070; |
| PR | 28-JUN-1999; | 99US-0140823; | PR | 13-SEP-1999; | 99US-0153758; |
| PR | 29-JUN-1999; | 99US-0140991; | PR | 15-SEP-1999; | 99US-0154018; |
| PR | 30-JUN-1999; | 99US-0141867; | PR | 16-SEP-1999; | 99US-0154039; |
| PR | 01-JUL-1999; | 99US-0141842; | PR | 20-SEP-1999; | 99US-0154779; |
| PR | 02-JUL-1999; | 99US-0142055; | PR | 22-SEP-1999; | 99US-0155133; |
| PR | 06-JUL-1999; | 99US-0142390; | PR | 23-SEP-1999; | 99US-0155133; |
| PR | 09-JUL-1999; | 99US-0142803; | PR | 24-SEP-1999; | 99US-0155486; |
| PR | 09-JUL-1999; | 99US-0142820; | PR | 28-SEP-1999; | 99US-0156458; |
| PR | 12-JUL-1999; | 99US-0142977; | PR | 29-SEP-1999; | 99US-01 |

| | | | | |
|-----------------------|----------------|-----------------|----------|------------|
| Query Match | 59.64 | Score 1052 | DB 21 | Length 239 |
| Best Local Similarity | 86.24 | Pred. No. 4e-82 | | |
| Matches 206 | Conservative 8 | Mismatches 21 | Indels 4 | Gaps 3 |

```

RESULT 7
AAG15886
ID      AAG15886 standard; Protein; 239 AA
XX

```

| | | |
|----|--------------|--------------|
| PR | 23-APR-1999; | 99US-0130891 |
| PR | 26-APR-1999; | 99US-0131449 |
| PR | 30-APR-1999; | 99US-0132048 |
| PR | 03-MAY-1999; | 99US-0132407 |
| PR | 04-MAY-1999; | 99US-0132484 |
| PR | 05-MAY-1999; | 99US-0132485 |
| PR | 06-MAY-1999; | 99US-0132486 |
| PR | 07-MAY-1999; | 99US-0132487 |
| PR | 11-MAY-1999; | 99US-0134256 |
| PR | 14-MAY-1999; | 99US-0134718 |
| PR | 14-MAY-1999; | 99US-0134719 |
| PR | 14-MAY-1999; | 99US-0134721 |
| PR | 18-MAY-1999; | 99US-0134768 |
| PR | 19-MAY-1999; | 99US-0135491 |
| PR | 21-MAY-1999; | 99US-0135524 |
| PR | 21-MAY-1999; | 99US-0135353 |
| PR | 25-MAY-1999; | 99US-0135629 |
| PR | 25-MAY-1999; | 99US-0136021 |
| PR | 28-MAY-1999; | 99US-0136392 |
| PR | 28-MAY-1999; | 99US-0136782 |
| PR | 01-JUN-1999; | 99US-0137222 |
| PR | 03-JUN-1999; | 99US-0137528 |
| PR | 04-JUN-1999; | 99US-0137502 |
| PR | 07-JUN-1999; | 99US-0137724 |
| PR | 08-JUN-1999; | 99US-0138094 |
| PR | 10-JUN-1999; | 99US-0138540 |
| PR | 14-JUN-1999; | 99US-0138647 |
| PR | 16-JUN-1999; | 99US-0138111 |
| PR | 16-JUN-1999; | 99US-0139452 |
| PR | 17-JUN-1999; | 99US-0139453 |
| PR | 18-JUN-1999; | 99US-0139454 |
| PR | 18-JUN-1999; | 99US-0139455 |
| PR | 18-JUN-1999; | 99US-0139456 |
| PR | 18-JUN-1999; | 99US-0139457 |
| PR | 18-JUN-1999; | 99US-0139458 |
| PR | 18-JUN-1999; | 99US-0139459 |
| PR | 18-JUN-1999; | 99US-0139460 |
| PR | 18-JUN-1999; | 99US-0139461 |
| PR | 18-JUN-1999; | 99US-0139462 |
| PR | 18-JUN-1999; | 99US-0139463 |
| PR | 18-JUN-1999; | 99US-0139750 |
| PR | 21-JUN-1999; | 99US-0139763 |
| PR | 21-JUN-1999; | 99US-0139817 |
| PR | 22-JUN-1999; | 99US-0139899 |
| PR | 23-JUN-1999; | 99US-0140353 |
| PR | 23-JUN-1999; | 99US-0140354 |
| PR | 24-JUN-1999; | 99US-0140695 |
| PR | 28-JUN-1999; | 99US-0140823 |
| PR | 30-JUN-1999; | 99US-0140981 |
| PR | 01-JUL-1999; | 99US-0141287 |
| PR | 01-JUL-1999; | 99US-0141842 |
| PR | 02-JUL-1999; | 99US-0142154 |
| PR | 06-JUL-1999; | 99US-0142055 |
| PR | 06-JUL-1999; | 99US-0142390 |
| PR | 08-JUL-1999; | 99US-0142803 |
| PR | 09-JUL-1999; | 99US-0142820 |
| PR | 12-JUL-1999; | 99US-0142977 |
| PR | 13-JUL-1999; | 99US-0143342 |
| PR | 15-JUL-1999; | 99US-0143624 |
| PR | 15-JUL-1999; | 99US-0144005 |
| PR | 16-JUL-1999; | 99US-0144085 |
| PR | 16-JUL-1999; | 99US-0144086 |
| PR | 19-JUL-1999; | 99US-0144325 |
| PR | 19-JUL-1999; | 99US-0144331 |
| PR | 19-JUL-1999; | 99US-0144332 |
| PR | 19-JUL-1999; | 99US-0144333 |
| PR | 19-JUL-1999; | 99US-0144334 |
| PR | 19-JUL-1999; | 99US-0144335 |
| PR | 20-JUL-1999; | 99US-0144352 |
| PR | 20-JUL-1999; | 99US-0144632 |

| Accession | Query Match | Score | DB | Length | Matches | Local Similarity | Conservative | Mismatches | Indels | Gaps |
|----------------|---|-------|------|--------|---------|------------------|--------------|------------|--------|------|
| PR 20-JUL-1999 | 98 IAGSRVTPYSSONQPLSSAFQSPYQVSPSSSPSPSPR-CEPNNNSTFPPL | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 1 MAGSSRPATYSSHNSQPLSTDSPTLISYQSSSSPSPSPRVDPHN--ISTIFPL | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 157 RNCGISPLSLRISNSCPVTPVSSPTSKNPPLPNWESIAKQSMAL-AKQSMASFPY | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 59 RNCGISPLSLRISNSCPVTPVSSPTSKNPPLPNWESIAKQSMAL-AKQSMASFPY | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 216 FYAVSAPASPTHRHQTTLATIEECBESDSSTVDSGHMISFQFAQQQPSASMTSP | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 119 FYAVSAPASPTHRHQTTLATIEECBESDSSTVDSGHMISFQFAQQQPSASMTSP | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 276 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86.24 | 8 | 21 | 4 | 3 |
| PR 21-JUL-1999 | 179 FNLVKPAPQMSPTTAFOEIGQSSSEKFEKNSOVKPMGEGRIDHVGMEDELTLGNGKA | 59.64 | 1052 | 239 | 206 | 86. | | | | |

PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135122.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140659.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151103.
PR 31-AUG-1999; 99US-0151138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.

PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 56.3%; Score 994; DB 21; Length 189;
Basic Local Similarity 99.5%; Pred. No. 2, 8e-77;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 148 MSSTPFPPLRNGIPSSLPRLRISNSCPYTPVSSPTSKPKPLPWESTIAKOSMAIAQ 207
1 MSTPFPPLRNGIPSSLPRLRISNSCPYTPVSSPTSKPKPLPWESTIAKOSMAIAQ 60
DB 1
OY 208 SMASFPYPRVAVSAPASPTHRHOFHTLATIPEDCEDSSSVDSGHWISFQKFAQQPFSA 267
61 SMASFPYPRVAVSAPASPTHRHOFHTPATIPEDCEDSSSVDSGHWISFQKFAQQPFSA 120
OY 266 SWVPPSPFTNLVVPARQOKSPNTAAROEIGQSEPKFENSQVPRWEGERIHDVGMEDLEL 327
121 SWVPPSPFTNLVVPARQOKSPNTAAROEIGQSEPKFENSQVPRWEGERIHDVGMEDLEL 180
DB 1
OY 328 TLGNKGARG 336
181 TLGNKGARG 189
DB 1

RESULT 9
AAG30139
ID AAG30139 standard, Protein; 284 AA.
XX AAG30139;
AC AAG30139;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 35979.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35979.
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN RP1033405-A2.
XX RP1033405-A2.
PD 06-SEP-2000.
XX 06-SEP-2000.
PR 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138547.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 26-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

```

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151348.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

```

```

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159296.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 30.7%; Score 541; DB 21; Length 284;
Best Local Similarity 40.2%; Pred. No. 3,8e-18;
Matches 143; Conservative 29; Mismatches 64; Indels 120; Gaps 17;

```

```

QY 23 RKQSWRENNRRERRRAVAKIYTGIRAGQYNI.PKHCDNNEVLCALCWEAGVVE 82
DB 6 RLPWKERENKRRERRRAIAKIIPAGLRIGHFKLPKHCDNNEVLCALCWEAGVVE 65
QY 83 DGTTRKCKPLPELAGTSSRVTYS-SQNSPLSASFQSPISYQVSPSSSEFSPSR 141
DB 66 DGTTRKCKPLPELAGTSSRVTYS-SQNSPLSASFQSPISYQVSPSSSEFSPSR 117
QY 142 --GEPNNMSSTFFPIARN--GGIPSLPLSR-LSNCCPTVPYSSPTSKNPKLPNWS 196
DB 118 PFGDAN-----SLIPWLKLNSSNSPSKLPFHGNSISAPVTPPLA----- 157
QY 197 IAKQSMALAKQSMASFVYFPAVSAPASPTRRHQFHLATIPCEDSDSTVDGHWISF 256
DB 158 -----RSPF-RDQ-----VILP-----DSG-WLS- 174
QY 257 OKFAQQQPSASAVPTSPFNLY-----KPA-POOM 286
DB 175 ---GKQTPQSG---PSSPFLSVSRNPFQXAFKMDSCNSPMKTPQSGNCSFAIPGV 228
QY 287 SPYTAAPQELIGSSSEPKF-----ENSQVQWEGERTH-DVGKEDLELTGNGAR 335
DB 229 DQNSDVPMADGWTAEPAFCGNMAAANGVYKPMWEGERTHGCVSDDLELTGNSRTR 284

```

```

RESULT 10
AAGS2728
ID AAGS2728 strand; Protein; 284 AA.
XX AAGS2728;
AC AAGS2728;
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana [protein fragment SEQ ID NO: 67061.
DE Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW

```

XX termination sequence.
XX Arabidopsis thaliana.
XX BP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137232.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149922.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150584.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 30.74; Score 541; DB 21; Length 284;
Best Local Similarity 40.23; Pred. No. 3.8e-38;
Matches 143; Conservative 29; Mismatches 64; Indels 120; Gaps 17;

QY 23 RKDSWREKNNRRERRRBAVAKIYTGRAQGDYMLPKGCDNNEVLKALCWEAGWVEE 82
DB 6 RTTWMKRENNKRRERRRRAIAKIFAGLRHGNFKLPKHCDDNNEVLKALCWEAGWVEE 65
QY 83 DGTTRKGCPLPGEIAGTSSRVTPYS-SQNSQPLSAPFQSPISQVSPSSSFPSPSR 141
DB 66 DGTTRKGCPLPGEIAGTSSRVTPYS-SQNSQPLSAPFQSPISQVSPSSSFPSPSR 141
QY 142 --GEPNNNSSTFFPLRN--GGIPSLPLSR--ISNSCPVTPPVSSPTSKNPKPLPWMS 196
DB 118 PEGDAN-----SLIPMLKNLSNSPSKLPFHGNSIAPTPPLA----- 157
QY 197 IAKOSMAIAKOSMAFVFPYAVASAPASPTHHQHTLATIPCEDESDSTVDSCHWISF 256
DB 158 -----RSPTR-RDQ-----VRIIP-----DSG-WLS- 174
QY 257 OKPAQOQPFASAVPTSPTENLV-----KPA-PQOM 286
DB 175 ---GNGTPOSG---PSSPTSLVSKNPFPEKAFKNGDCNSPMWTPQSGNCSPPAIPAGV 228
QY 287 SPNTAAPFOEIGSSSEKPF-----ENSQVPMWGERIH-DVGMEDLELTUNGKAR 335
DB 229 DONSVDPMADGMWTAEPACGNAMAMANGMVPWMEGERIHCEVCSDDELTUNGSRTR 284

RESULT 11
ID AAG29487 standard; Protein; 169 AA.
XX AAG29487;
AC AAG29487;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment 600 ID NO: 35094.
DE Arabidopsis thaliana protein fragment 600 ID NO: 35094.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134371.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140595.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161995.
PR 29-OCT-1999; 99US-0162142.

Query Match 18.7%; Score 310; DB 21; Length 169;
Best Local Similarity 46.3%; Pred. No. 2,4e-20;
Matches 68; Conservative 23; Mismatches 48; Indels 8; Gaps 3;

QY 10 SAAAAAARRRRRRRRRRRRRRRAVAAKTYTGIRAGGDYTLPRGCDNNEVL 69
DB 3 AGGGGGGGSSGRPTWKERENNNKKRRRRRAIRAKTYSGIRAGGNYTLPRGCDNNEVL 62
QY 70 KALCVBAGWVVEBDGTTTRKCK--PLPGETAGTSSVTTPVSSONOSPLSAFOSPIRS 126
DB 63 KALCLEAGWIVEDDITTRKGFSGHOHOLFEDLQTSAGI--HOSNOVHNHOLFQ--VJH 117
QY 127 YOVSPSSSSFPSPRSGEBNNNNMSTFF 153
DB 118 LRTTEVOHHFBRVHLAMTETLHTTFF 144

RESULT 12
AAG30141
ID AAG30141 standard; Protein; 205 AA.
XX
AC AAG30141;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35981.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000, 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128774.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 16-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137722.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

```

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151005.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152353.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154079.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 16-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13 34; Score 235.5; DB 21; Length 205;
Best Local Similarity 32.14; Pred. No. 4, 1e-12;
Matches 86; Conservative 18; Mismatches 53; Indels 111; Gaps 15;

```

```

DB 81 PT-RDQ-----VTPP-----DSG-WLS-----GMOTPOSG---DSPTFLVSRNPF 117
QY 280 -----KPA-POOMSPNTAFOEIGOSSEFKF-----ENSQ 308
DB 118 FDKEAFKMGDCNSPMTTPOGSGNCSPFAPAGVQONSUVPAADGNTAFATGCMMAANGM 177
QY 309 VKPWEGERIH-DVGMEDLETLTGNGKAR 335
DB 178 VKPWEGERIHGCVSDDELTLGNSRR 205

RESULT 13
AGS2730
ID AGS2730 standard; Protein; 205 AA.
AC
AC AGS2730;
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67063.
XX
XX Protein identification, signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123160.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135124.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.

```

| | | | | | |
|-------|--------------|--|-----|-----------------------|---------------|
| PR | 10-JUN-1999; | 99US-0138540. | PR | 09-AUG-1999; | 99US-0147935. |
| PR | 10-JUN-1999; | 99US-0138847. | PR | 10-AUG-1999; | 99US-0148171. |
| PR | 14-JUN-1999; | 99US-0139119. | PR | 11-AUG-1999; | 99US-0148319. |
| PR | 16-JUN-1999; | 99US-0139452. | PR | 12-AUG-1999; | 99US-0148341. |
| PR | 16-JUN-1999; | 99US-0139453. | PR | 13-AUG-1999; | 99US-0148565. |
| PR | 17-JUN-1999; | 99US-0139492. | PR | 13-AUG-1999; | 99US-0148565. |
| PR | 18-JUN-1999; | 99US-0139454. | PR | 16-AUG-1999; | 99US-0148684. |
| PR | 18-JUN-1999; | 99US-0139455. | PR | 17-AUG-1999; | 99US-0149368. |
| PR | 18-JUN-1999; | 99US-0139456. | PR | 18-AUG-1999; | 99US-0149175. |
| PR | 18-JUN-1999; | 99US-0139457. | PR | 20-AUG-1999; | 99US-0149722. |
| PR | 18-JUN-1999; | 99US-0139458. | PR | 20-AUG-1999; | 99US-0149723. |
| PR | 18-JUN-1999; | 99US-0139459. | PR | 20-AUG-1999; | 99US-0149929. |
| PR | 18-JUN-1999; | 99US-0139460. | PR | 23-AUG-1999; | 99US-0149902. |
| PR | 18-JUN-1999; | 99US-0139461. | PR | 23-AUG-1999; | 99US-0149930. |
| PR | 18-JUN-1999; | 99US-0139462. | PR | 25-AUG-1999; | 99US-0150566. |
| PR | 18-JUN-1999; | 99US-0139463. | PR | 26-AUG-1999; | 99US-0150884. |
| PR | 18-JUN-1999; | 99US-0139750. | PR | 27-AUG-1999; | 99US-0151065. |
| PR | 21-JUN-1999; | 99US-0139763. | PR | 27-AUG-1999; | 99US-0151066. |
| PR | 22-JUN-1999; | 99US-0139817. | PR | 27-AUG-1999; | 99US-0151080. |
| PR | 23-JUN-1999; | 99US-0140353. | PR | 30-AUG-1999; | 99US-0151303. |
| PR | 23-JUN-1999; | 99US-0140354. | PR | 31-AUG-1999; | 99US-0151438. |
| PR | 24-JUN-1999; | 99US-0140695. | PR | 01-SEP-1999; | 99US-0151930. |
| PR | 28-JUN-1999; | 99US-0140823. | PR | 07-SEP-1999; | 99US-0152363. |
| PR | 29-JUN-1999; | 99US-0140991. | PR | 10-SEP-1999; | 99US-0153763. |
| PR | 30-JUN-1999; | 99US-0141287. | PR | 13-SEP-1999; | 99US-0153758. |
| PR | 01-JUL-1999; | 99US-0141842. | PR | 15-SEP-1999; | 99US-0154018. |
| PR | 02-JUL-1999; | 99US-0141544. | PR | 16-SEP-1999; | 99US-0154039. |
| PR | 06-JUL-1999; | 99US-0142055. | PR | 20-SEP-1999; | 99US-0154779. |
| PR | 08-JUL-1999; | 99US-0142390. | PR | 22-SEP-1999; | 99US-0155139. |
| PR | 09-JUL-1999; | 99US-0142803. | PR | 23-SEP-1999; | 99US-0155486. |
| PR | 12-JUL-1999; | 99US-0142920. | PR | 24-SEP-1999; | 99US-0155659. |
| PR | 13-JUL-1999; | 99US-0142977. | PR | 28-SEP-1999; | 99US-0156458. |
| PR | 14-JUL-1999; | 99US-0143342. | PR | 29-SEP-1999; | 99US-0156596. |
| PR | 15-JUL-1999; | 99US-0143624. | PR | 04-OCT-1999; | 99US-0157117. |
| PR | 16-JUL-1999; | 99US-0144005. | PR | 05-OCT-1999; | 99US-0157753. |
| PR | 16-JUL-1999; | 99US-0144085. | PR | 06-OCT-1999; | 99US-0157865. |
| PR | 19-JUL-1999; | 99US-0144322. | PR | 07-OCT-1999; | 99US-0158029. |
| PR | 19-JUL-1999; | 99US-0144331. | PR | 12-OCT-1999; | 99US-0158232. |
| PR | 19-JUL-1999; | 99US-0144332. | PR | 13-OCT-1999; | 99US-0158369. |
| PR | 19-JUL-1999; | 99US-0144333. | PR | 13-OCT-1999; | 99US-0159293. |
| PR | 19-JUL-1999; | 99US-0144334. | PR | 13-OCT-1999; | 99US-0159294. |
| PR | 19-JUL-1999; | 99US-0144335. | PR | 14-OCT-1999; | 99US-0159329. |
| PR | 20-JUL-1999; | 99US-0144352. | PR | 14-OCT-1999; | 99US-0159330. |
| PR | 20-JUL-1999; | 99US-0144632. | PR | 14-OCT-1999; | 99US-0159331. |
| PR | 21-JUL-1999; | 99US-0144814. | PR | 14-OCT-1999; | 99US-0159637. |
| PR | 21-JUL-1999; | 99US-0145086. | PR | 18-OCT-1999; | 99US-0159638. |
| PR | 21-JUL-1999; | 99US-0145088. | PR | 21-OCT-1999; | 99US-0159584. |
| PR | 22-JUL-1999; | 99US-0145085. | PR | 21-OCT-1999; | 99US-0160741. |
| PR | 22-JUL-1999; | 99US-0145087. | PR | 21-OCT-1999; | 99US-0160767. |
| PR | 22-JUL-1999; | 99US-0145089. | PR | 21-OCT-1999; | 99US-0160768. |
| PR | 23-JUL-1999; | 99US-0145129. | PR | 21-OCT-1999; | 99US-0160770. |
| PR | 23-JUL-1999; | 99US-0145145. | PR | 21-OCT-1999; | 99US-0160814. |
| PR | 23-JUL-1999; | 99US-0145218. | PR | 22-OCT-1999; | 99US-0160815. |
| PR | 26-JUL-1999; | 99US-0145224. | PR | 22-OCT-1999; | 99US-0160980. |
| PR | 27-JUL-1999; | 99US-0145276. | PR | 22-OCT-1999; | 99US-0160981. |
| PR | 27-JUL-1999; | 99US-0145293. | PR | 25-OCT-1999; | 99US-0160989. |
| PR | 27-JUL-1999; | 99US-0145918. | PR | 25-OCT-1999; | 99US-0161404. |
| PR | 28-JUL-1999; | 99US-0145919. | PR | 25-OCT-1999; | 99US-0161405. |
| PR | 02-AUG-1999; | 99US-0145951. | PR | 26-OCT-1999; | 99US-0161359. |
| PR | 02-AUG-1999; | 99US-0146386. | PR | 26-OCT-1999; | 99US-0161360. |
| PR | 03-AUG-1999; | 99US-0146388. | PR | 26-OCT-1999; | 99US-0161361. |
| PR | 03-AUG-1999; | 99US-0146389. | PR | 28-OCT-1999; | 99US-0161920. |
| PR | 04-AUG-1999; | 99US-0147038. | PR | 28-OCT-1999; | 99US-0161922. |
| PR | 04-AUG-1999; | 99US-0147204. | PR | 28-OCT-1999; | 99US-0161993. |
| PR | 05-AUG-1999; | 99US-0147302. | PR | 29-OCT-1999; | 99US-0162142. |
| PR | 06-AUG-1999; | 99US-0147260. | | | |
| PR | 06-AUG-1999; | 99US-0147303. | | | |
| PR | 06-AUG-1999; | 99US-0147416. | | | |
| PR | 09-AUG-1999; | 99US-0147493. | | | |
| <hr/> | | | | | |
| PR | 09-AUG-1999; | 99US-0147935. | PR | 09-AUG-1999; | 99US-0147935. |
| PR | 10-AUG-1999; | 99US-0148171. | PR | 10-AUG-1999; | 99US-0148171. |
| PR | 11-AUG-1999; | 99US-0148319. | PR | 11-AUG-1999; | 99US-0148319. |
| PR | 12-AUG-1999; | 99US-0148341. | PR | 12-AUG-1999; | 99US-0148341. |
| PR | 13-AUG-1999; | 99US-0148565. | PR | 13-AUG-1999; | 99US-0148565. |
| PR | 13-AUG-1999; | 99US-0148565. | PR | 13-AUG-1999; | 99US-0148565. |
| PR | 16-AUG-1999; | 99US-0148684. | PR | 16-AUG-1999; | 99US-0148684. |
| PR | 17-AUG-1999; | 99US-0149368. | PR | 17-AUG-1999; | 99US-0149368. |
| PR | 18-AUG-1999; | 99US-0149175. | PR | 18-AUG-1999; | 99US-0149175. |
| PR | 20-AUG-1999; | 99US-0149722. | PR | 20-AUG-1999; | 99US-0149722. |
| PR | 20-AUG-1999; | 99US-0149723. | PR | 20-AUG-1999; | 99US-0149723. |
| PR | 20-AUG-1999; | 99US-0149929. | PR | 20-AUG-1999; | 99US-0149929. |
| PR | 23-AUG-1999; | 99US-0149902. | PR | 23-AUG-1999; | 99US-0149902. |
| PR | 23-AUG-1999; | 99US-0149930. | PR | 23-AUG-1999; | 99US-0149930. |
| PR | 25-AUG-1999; | 99US-0150566. | PR | 25-AUG-1999; | 99US-0150566. |
| PR | 26-AUG-1999; | 99US-0150884. | PR | 26-AUG-1999; | 99US-0150884. |
| PR | 27-AUG-1999; | 99US-0151065. | PR | 27-AUG-1999; | 99US-0151065. |
| PR | 27-AUG-1999; | 99US-0151066. | PR | 27-AUG-1999; | 99US-0151066. |
| PR | 30-AUG-1999; | 99US-0151080. | PR | 30-AUG-1999; | 99US-0151080. |
| PR | 31-AUG-1999; | 99US-0151438. | PR | 31-AUG-1999; | 99US-0151438. |
| PR | 01-SEP-1999; | 99US-0151930. | PR | 01-SEP-1999; | 99US-0151930. |
| PR | 07-SEP-1999; | 99US-0152363. | PR | 07-SEP-1999; | 99US-0152363. |
| PR | 10-SEP-1999; | 99US-0153763. | PR | 10-SEP-1999; | 99US-0153763. |
| PR | 13-SEP-1999; | 99US-0153758. | PR | 13-SEP-1999; | 99US-0153758. |
| PR | 15-SEP-1999; | 99US-0154018. | PR | 15-SEP-1999; | 99US-0154018. |
| PR | 16-SEP-1999; | 99US-0154039. | PR | 16-SEP-1999; | 99US-0154039. |
| PR | 20-SEP-1999; | 99US-0154779. | PR | 20-SEP-1999; | 99US-0154779. |
| PR | 22-SEP-1999; | 99US-0155139. | PR | 22-SEP-1999; | 99US-0155139. |
| PR | 23-SEP-1999; | 99US-0155486. | PR | 23-SEP-1999; | 99US-0155486. |
| PR | 24-SEP-1999; | 99US-0155659. | PR | 24-SEP-1999; | 99US-0155659. |
| PR | 28-SEP-1999; | 99US-0156458. | PR | 28-SEP-1999; | 99US-0156458. |
| PR | 29-SEP-1999; | 99US-0156596. | PR | 29-SEP-1999; | 99US-0156596. |
| PR | 04-OCT-1999; | 99US-0157117. | PR | 04-OCT-1999; | 99US-0157117. |
| PR | 05-OCT-1999; | 99US-0157753. | PR | 05-OCT-1999; | 99US-0157753. |
| PR | 06-OCT-1999; | 99US-0157865. | PR | 06-OCT-1999; | 99US-0157865. |
| PR | 07-OCT-1999; | 99US-0158029. | PR | 07-OCT-1999; | 99US-0158029. |
| PR | 12-OCT-1999; | 99US-0158232. | PR | 12-OCT-1999; | 99US-0158232. |
| PR | 13-OCT-1999; | 99US-0158369. | PR | 13-OCT-1999; | 99US-0158369. |
| PR | 13-OCT-1999; | 99US-0159293. | PR | 13-OCT-1999; | 99US-0159293. |
| PR | 13-OCT-1999; | 99US-0159294. | PR | 13-OCT-1999; | 99US-0159294. |
| PR | 14-OCT-1999; | 99US-0159329. | PR | 14-OCT-1999; | 99US-0159329. |
| PR | 14-OCT-1999; | 99US-0159330. | PR | 14-OCT-1999; | 99US-0159330. |
| PR | 14-OCT-1999; | 99US-0159331. | PR | 14-OCT-1999; | 99US-0159331. |
| PR | 14-OCT-1999; | 99US-0159637. | PR | 14-OCT-1999; | 99US-0159637. |
| PR | 18-OCT-1999; | 99US-0159638. | PR | 18-OCT-1999; | 99US-0159638. |
| PR | 21-OCT-1999; | 99US-0159584. | PR | 21-OCT-1999; | 99US-0159584. |
| PR | 21-OCT-1999; | 99US-0160741. | PR | 21-OCT-1999; | 99US-0160741. |
| PR | 21-OCT-1999; | 99US-0160767. | PR | 21-OCT-1999; | 99US-0160767. |
| PR | 21-OCT-1999; | 99US-0160768. | PR | 21-OCT-1999; | 99US-0160768. |
| PR | 21-OCT-1999; | 99US-0160770. | PR | 21-OCT-1999; | 99US-0160770. |
| PR | 21-OCT-1999; | 99US-0160814. | PR | 21-OCT-1999; | 99US-0160814. |
| PR | 22-OCT-1999; | 99US-0160815. | PR | 22-OCT-1999; | 99US-0160815. |
| PR | 22-OCT-1999; | 99US-0160980. | PR | 22-OCT-1999; | 99US-0160980. |
| PR | 22-OCT-1999; | 99US-0160981. | PR | 22-OCT-1999; | 99US-0160981. |
| PR | 25-OCT-1999; | 99US-0160989. | PR | 25-OCT-1999; | 99US-0160989. |
| PR | 25-OCT-1999; | 99US-0161404. | PR | 25-OCT-1999; | 99US-0161404. |
| PR | 25-OCT-1999; | 99US-0161405. | PR | 25-OCT-1999; | 99US-0161405. |
| PR | 26-OCT-1999; | 99US-0161359. | PR | 26-OCT-1999; | 99US-0161359. |
| PR | 26-OCT-1999; | 99US-0161360. | PR | 26-OCT-1999; | 99US-0161360. |
| PR | 26-OCT-1999; | 99US-0161361. | PR | 26-OCT-1999; | 99US-0161361. |
| PR | 28-OCT-1999; | 99US-0161920. | PR | 28-OCT-1999; | 99US-0161920. |
| PR | 28-OCT-1999; | 99US-0161922. | PR | 28-OCT-1999; | 99US-0161922. |
| PR | 28-OCT-1999; | 99US-0161993. | PR | 28-OCT-1999; | 99US-0161993. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 29-OCT-1999; | 99US-0162142. |
| <hr/> | | | | | |
| QY | 110 | SONQSPUSAFQSPRISYQVSPSSSPSPSR--GEINNMSTPFPFLN--GGIPSSL | 165 | Query Match | 113.3%; |
| | | | | Best Local Similarity | 32.1%; |
| | | | | Matches | 86; |
| | | | | Conservative | 18; |
| | | | | Mismatches | 53; |
| | | | | Indels | 111; |
| | | | | Gaps | 15; |


```
DB      7 STSASPCSSYQHSPPRASINPSSSSPPTNPFCDAN-----SLPWLKNLSNBSKL 61
OY      166 PSLR-1SNSCPVTTPVSPPTSKNPKPLPMWESIAKOSMAIAKOSMASFNYPYAVSAPAS 224
DB      62 PPHNGISIAPIVPPPLA-----RS 80
OY      225 PTHRHQPHITATITPECDSDSTVDSGHWISFQKPAQOPFSASMTPTSTFNLV----- 279
DB      81 PT-RDQ-----VTP-----DSG-WLS-----GMQTPQSG-----PSSPTFLVSRNP 117
OY      280 -----KPA-POQMSPTAFAEIQSSSEKFR-----ENSG 308
DB      118 FDXEAFKMGDCNSPMWTPGSGNCSPAI PAVGDQNSDVPMADGNTAFAGCNMAANGM 177
OY      309 VKPMEGERIIV-DVGMEDLETLGNGKAR 335
DB      178 VKPMEGERIHGEVSDLETLGNSRTR 205

RESULT 14
AAG30140
ID      AAG30140 standard; Protein; 208 AA.
XX      AAG30140;
XX      17-OCT-2000 (first entry)
XX      Arabidopsis thaliana protein fragment SEQ ID NO: 35980.
XX      Protein identification; signal transduction pathway; metabolic pathway;
XX      hybridisation assay; genetic mapping; gene expression control; promoter;
XX      termination sequence.
XX      Arabidopsis thaliana.
XX      EPI033405-A2.
XX      06-SEP-2000.
XX      25-FEB-2000; 2000EP-0301439.
XX      25-FEB-1999; 99US-0121825.
XX      05-MAR-1999; 99US-0123180.
XX      09-MAR-1999; 99US-0123548.
XX      23-MAR-1999; 99US-0125788.
XX      25-MAR-1999; 99US-0126264.
XX      29-MAR-1999; 99US-0126785.
XX      01-APR-1999; 99US-0127462.
XX      06-APR-1999; 99US-0128234.
XX      08-APR-1999; 99US-0128714.
XX      15-APR-1999; 99US-0129845.
XX      19-APR-1999; 99US-0130077.
XX      21-APR-1999; 99US-0130449.
XX      23-APR-1999; 99US-0130510.
XX      28-APR-1999; 99US-0130891.
XX      30-APR-1999; 99US-0131449.
XX      30-APR-1999; 99US-0132048.
XX      30-APR-1999; 99US-0132407.
XX      04-MAY-1999; 99US-0132484.
XX      05-MAY-1999; 99US-0132485.
XX      06-MAY-1999; 99US-0132486.
XX      06-MAY-1999; 99US-0132487.
XX      07-MAY-1999; 99US-0132863.
XX      11-MAY-1999; 99US-0134256.
XX      14-MAY-1999; 99US-0134218.
XX      14-MAY-1999; 99US-0134219.
XX      14-MAY-1999; 99US-0134321.
XX      14-MAY-1999; 99US-0134370.
XX      18-MAY-1999; 99US-0134768.
XX      19-MAY-1999; 99US-0134941.
XX      20-MAY-1999; 99US-0135124.
XX      21-MAY-1999; 99US-0135353.
```

```
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139457.
PR      17-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140685.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      02-JUL-1999; 99US-0142154.
PR      06-JUL-1999; 99US-0142055.
PR      08-JUL-1999; 99US-0142380.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145145.
PR      22-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145224.
PR      23-JUL-1999; 99US-0145228.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
```

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147192.
PR 07-AUG-1999; 99US-0147260.
PR 08-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 10-AUG-1999; 99US-0147493.
PR 11-AUG-1999; 99US-0148177.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 14-AUG-1999; 99US-0148684.
PR 15-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149722.
PR 19-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 21-AUG-1999; 99US-0149902.
PR 22-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 24-AUG-1999; 99US-0150884.
PR 25-AUG-1999; 99US-0151065.
PR 26-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 28-AUG-1999; 99US-0151303.
PR 29-AUG-1999; 99US-0151438.
PR 30-AUG-1999; 99US-0151930.
PR 31-AUG-1999; 99US-0152363.
PR 01-SEP-1999; 99US-0153070.
PR 02-SEP-1999; 99US-0153758.
PR 03-SEP-1999; 99US-0154018.
PR 04-SEP-1999; 99US-0154039.
PR 05-SEP-1999; 99US-0154779.
PR 06-SEP-1999; 99US-0155139.
PR 07-SEP-1999; 99US-0155486.
PR 08-SEP-1999; 99US-0155659.
PR 09-SEP-1999; 99US-0156458.
PR 10-SEP-1999; 99US-0156596.
PR 11-SEP-1999; 99US-0157753.
PR 12-SEP-1999; 99US-0157865.
PR 13-SEP-1999; 99US-0158029.
PR 14-SEP-1999; 99US-0158232.
PR 15-SEP-1999; 99US-0158369.
PR 16-SEP-1999; 99US-0159293.
PR 17-SEP-1999; 99US-0159294.
PR 18-SEP-1999; 99US-0159295.
PR 19-SEP-1999; 99US-0159329.
PR 20-SEP-1999; 99US-0159330.
PR 21-SEP-1999; 99US-0159331.
PR 22-SEP-1999; 99US-0159637.
PR 23-SEP-1999; 99US-0159638.
PR 24-SEP-1999; 99US-015984.
PR 25-SEP-1999; 99US-0160741.
PR 26-SEP-1999; 99US-0160767.
PR 27-SEP-1999; 99US-0160768.
PR 28-SEP-1999; 99US-0160770.
PR 29-SEP-1999; 99US-0160814.
PR 30-SEP-1999; 99US-0160815.
PR 01-OCT-1999; 99US-0160980.
PR 02-OCT-1999; 99US-0160981.
PR 03-OCT-1999; 99US-0160989.
PR 04-OCT-1999; 99US-0161405.
PR 05-OCT-1999; 99US-0161406.
PR 06-OCT-1999; 99US-0161359.
PR 07-OCT-1999; 99US-0161360.
PR 08-OCT-1999; 99US-0161361.
PR 09-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0161993.
PR 30-OCT-1999; 99US-0162142.
Query Match 13.3%; Score 235.5; DB 21; Length 208;
Best Local Similarity 32.1%; Pred. No. 4, 2e-12;
Matches 86; Conservative 18; Mismatches 53; Indels 111; Gaps 15;
Db 110 SONOSPLSSAFOSPIPSYOVSPSSSPSPSR--GEPNNMSTFPFLRN--CGIPSL 165
10 STSAPSCSYQHPSPASINPSPSSSPPTFPQDAN-----SLIPWLNSSNSPSKL 64
166 PSLR-INSNCVTPPVSSPSPKPPPLPWNESIAKQSMIAKQSMASFPYFVAAGAPAS 224
65 PFFGNSTISAPVTPPLA-----RS 83
Qy 225 PTHRHQPTLATIPBCDSDSSTVDSCGWISQFAQQQPSASMPVPTNLV----- 279
84 PT-RDQ-----VTIP-----DSG-WLS---(MQTPQSG---PSSPFTSLVRNPF 120
Qy 280 -----KPA-PQMSPTVIAFOELGOSSEPKF-----ENQ 308
Db 121 FDKAFKMGDCNGSPMTWPGSCGNCSPALPAGVDQNDVPMADGWTAEFAFGCMAAANGM 180
Qy 309 VKPWEGERRIH-DYGMEDLELTGNGKAR 335
181 VKPWEGERRIHQEGVSDDELTLTGNSRR 208
RESULT 15
AAG52729
ID AAG52729 standard; Protein; 208 AA.
XX AAG52729;
AC AAG52729;
XX 18-OCT-2000 (first entry)
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 67062.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67062.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN EP1033405-A2.
FD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126285.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140625.
PR 28-JUN-1999; 99US-0140635.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150066.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159350.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159684.
PR 16-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 13.3% Score 235.5; DB 21; Length 208;
 Best Local Similarity 32.1%; Pred. No. 4,2e-12;
 Matches 86; Conservative 18; Mismatches 53; Indels 11; Gaps 15;

QY 110 SONQSPUSSAFQSPISYQVSPSSSPSPSR--GEPNNMSTFFPPLRN--GGIPSL 165
 DB 10 STSAPCSSYQHSPPRASYNPSPSSSPSPPTNPFCDAN---SLIPWLKULSSNSPSKL 64
 QY 166 PSUR-IGNSCVTPPVSSPTSKNPKPLPWNESIAKQSMALAKQSMASFNYPYAVASAPAS 224
 DB 65 PFFHNSISALVTPPLA-----RS 83
 QY 225 PTHRHOFHTLATIEPCDESDSSTVDSGHMSFOKFAQQQPFASNVPTPTNLV----- 279
 DB 84 PT-RDQ-----VTIP-----DSG-WLS---GMQTPOSG--PSSPTFSLVSRNP 120
 QY 280 -----KPA-PQOMSPNTAFAQEIQOSSEFKF-----ENSQ 308
 DB 121 FDKEAFKMGDCNSPMWTPGQSGNCSPALPAGVDONSVDVPMADGWTAEFAFGCNMAANCM 180
 QY 309 VKPWEGERRIH-DVGMEDLELTGNGKAR 335
 DB 181 VKPWEGERRIHGEVSDDELTGNSRTR 208

Search completed: December 19, 2003, 14:25:07
 Job time : 45 secs

QY 23 RKPMSRENNRRRRRAVAAKIYTGRLPAQDYNLPRKCDNNEVLKALCVRGAWVEE 82
 DB 6 RMPYWERENNNKRRRRRAIAAKIPTGLRMVGNVELPKKCDNNEVLKALCNRAGIYVER 65
 QY 83 DGTTRKGC-KPLPG-ELAGTSRVTPYSSONQSPLSAFOSPIPYQVSPSSSSPSPSP 140
 DB 66 DGTTRKGCSPRVERWEIGGSGATAPSCSSYOPSCASVYSPGSSNFMSPASSPANLT 125
 QY 141 RGEPPNNMSTFPPLRNGGIPSS---LGLRISN-----SCPYPVSPSPSKNP 188
 DB 126 SGD-----GSLIPWLKHLSTSSSSASSSRUPNLYIPGGSISAPVYPLSPARPP 180
 QY 189 KPLPMWESIAKQSMALAKQSMASFNYPYAVASAPASPTRHOPHTLATIPEDCESDSTV 248
 DB 181 RMTDWMQOL-----NNSFVVSSTPSPPTQ-----IIPDE----- 211
 QY 249 DSGHITSFQKPAQOQPFASMPVSTSTPVLVYKPAQOQSPNTAAFOE----- 295
 DB 212 ---WPSGIQLAQSS-----VPASPTPLV-----SQNPFGFEBAASAAGGGGGR 253
 QY 296 ---IGQSS-----EPKFE---ENSOVPMWEGRIH-DVQNE 323
 DB 254 MMTPEQSGTCSFPAIPPGADQTDADVPMSEAVAPBEFAFGSNTGLVYVMEGERIHESGSD 313
 QY 324 DLELTGNGKAR 335
 DB 314 DLELTGNSSTR 325

RESULT 5

T05026
 hypothetical protein P13C5_60 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear creas)
 C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
 C/Accession: T05026
 R/Beyan, M.; Pohl, T.; Weizenecker, T.; Bancroft, I.; Mewes, H.W.; Meyer, K.F.X.; Schuel
 submitted to the Protein Sequence Database, February 1998
 A/Reference number: Z15395
 A/Accession: T05026
 A/Molecule type: DNA
 A/Residues: 1-284 <BEV>
 A/Cross-references: EMBL:AL021711
 A/Experimental source: cultivar Columbia; BAC clone P13C5
 C/Genetics:
 A/Map position: 4
 A/Intron: 72/3
 A/Note: P13C5_60

Query Match 30.7%; Score 541; DB 2; Length 284;

Best Local Similarity 40.2%; Pred. No. 1.1e-29;

Matches 143; Conservative 29; Mismatches 64; Indels 120; Gaps 17;

QY 23 RKPMSRENNRRRRRAVAAKIYTGRLPAQDYNLPRKCDNNEVLKALCVRGAWVEE 82
 DB 6 RMPYWERENNNKRRRRRAIAAKIPTGLRMVGNVELPKKCDNNEVLKALCNRAGIYVER 65
 QY 83 DGTTRKGC-KPLPG-ELAGTSRVTPYSSONQSPLSAFOSPIPYQVSPSSSSPSPSP 141
 DB 66 DGTTRKGCSPRVERWEIGGSGATAPSCSSYOPSCASVYSPGSSNFMSPASSPANLT 125
 QY 141 RGEPPNNMSTFPPLRNGGIPSS---LGLRISN-----SCPYPVSPSPSKNP 188
 DB 126 SGD-----GSLIPWLKHLSTSSSSASSSRUPNLYIPGGSISAPVYPLSPARPP 180
 QY 189 KPLPMWESIAKQSMALAKQSMASFNYPYAVASAPASPTRHOPHTLATIPEDCESDSTV 248
 DB 181 RMTDWMQOL-----NNSFVVSSTPSPPTQ-----IIPDE----- 211
 QY 249 DSGHITSFQKPAQOQPFASMPVSTSTPVLVYKPAQOQSPNTAAFOE----- 295
 DB 212 ---WPSGIQLAQSS-----VPASPTPLV-----SQNPFGFEBAASAAGGGGGR 253
 QY 296 ---IGQSS-----EPKFE---ENSOVPMWEGRIH-DVQNE 323
 DB 254 MMTPEQSGTCSFPAIPPGADQTDADVPMSEAVAPBEFAFGSNTGLVYVMEGERIHESGSD 313
 QY 324 DLELTGNGKAR 335
 DB 314 DLELTGNSSTR 325

DB 229 DQNSDVPADGWTAEPAFCGNMAAAGVYPMWEGRIHGEVSDDELTGNSKTR 284

RESULT 6

T02459
 probable beta-amylase At2g45880 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear creas)
 C/Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 02-Feb-2001
 C/Accession: T02459; A04896
 R/Bounslay, S.D.; Lin, X.; Kechum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kau
 submitted to the EMBL Data Library, August 1998
 A/Description: Arabidopsis thaliana chromosome II BAC F4118 genomic sequence.
 A/Reference number: T02459
 A/Accession: T02459
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-687 <ROU>
 A/Cross-references: EMBL:AC004665; NID:g3386593; PID:g3386606
 A/Experimental source: cultivar Columbia
 R/Lin, X.; Kaul, S.; Bounslay, S.D.; Shea, T.P.; Genito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Xoo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.W.; Venter,
 Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A04420
 A/Accession: A04896
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-687 <STO>
 A/Cross-references: GB:AB002093; NID:g3386606; PID:AA028536.1; GSPDB:GN00139
 C/Genetics:
 A/Map position: 2
 A/Introns: 138/3; 229/3; 255/3; 323/3; 388/3; 443/3; 530/3; 603/3

Query Match 11.5%; Score 203.5; DB 2; Length 687;
 Best Local Similarity 39.5%; Pred. No. 2.3e-06;
 Matches 51; Conservative 16; Mismatches 53; Indels 9; Gaps 3;

QY 21 ARKPSRENNRRRRRAVAAKIYTGRLPAQDYNLPRKCDNNEVLKALCVRGAWV 80
 DB 67 SRRSPLEKERTKLRHRRRAITARIISGLRHHGNVYLRYRADINDVIALAREGWV 126

QY 81 EEDGTTY---RKGCPLRGEIA-GTSSRVTPYSSONQSP-----LSAFOSPIPYQVSP 131
 DB 127 LPDGTTFPSKSGRTKPTGSSAIVAAGSSASHISQCTSPALRVVSSGLRSPVELSSCRM 186

QY 132 SSSSPSPSPS 140
 DB 187 KGVFTPAPS 195

RESULT 7

A34092
 DNA-directed RNA polymerase (EC 2.7.7.6) II large chain - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: A34092
 R/Bird, D.M.; Riddle, D.L.
 Mol. Cell. Biol. 9, 4119-4130, 1989
 A/Title: Molecular cloning and sequencing of ama-1, the gene encoding the largest subunit
 A/Reference number: A34092; MUID:90066416; PMID:2566513
 A/Accession: A34092
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1859 <BIR>
 A/Cross-references: GB:M29295; NID:g156404; PID:AAA28126.1; PID:g156405
 A/Note: the authors failed to translated the codon GAG for residue 917 as Glu, and CAT f
 C/Superfamily: human DNA-directed RNA polymerase II large chain
 C/Keywords: DNA binding; nucleocytoplasmic transferase; tandem repeat; transcription; zinc finger
 Query Match 8.3%; Score 146; DB 1; Length 1859;

Db 237 PIPSTTSVQSSSSAPTTTSATSVQSSSSSTPIPTTSVQSSSSSAPTTSATSVQSS 296
 Qy 134 SAFP-----SPSAGEPNNMSTFFPLRNGGIP-SLPLRLRISNCEPTVPVSS-- 182
 Db 297 SSSPPLSSTTSVQSSSSSPTTSTTSVQSSSSGSAPTTSATSVQSSSS--SPPLSST 354
 Qy 183 ---PTSKNPKPLPMWESIAKQSMIAKQSMASFNYPFYAASAPLTHRQFTLTATPE 239
 Db 355 SVQSSSSSSPTTSTTSVQSSSSGSAPTTSATSVQSSSSSVPT-----TSATSVR 405
 Qy 240 CDESSSTVDSGHWISFQFAQQPFSASMWPTS-----PTFNLVLRPAQ 284
 Db 406 SSSSSSTPIPT-----TTSVQSSSSSVPTTSATSVQSSSSSTPIPTTSVQSSS 457
 Qy 285 QMSPTAA 292
 Db 458 SSAPTTSA 465

RESULT 11

S50755

hypothetical protein VSP-3 - Chlamydomonas reinhardtii
 C/Species: Chlamydomonas reinhardtii
 C/Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #ext_change 21-Jul-2000
 C/Accession: S50755

R/Moensner, J.P.; Molendijk, A.J.; van Egmond, P.; Klie, F.M.; Goodenough, U.W.; Harting, Plant Mol. Biol. 26, 947-960, 1994
 A/Title: Domain conservation in several volvocalean cell wall proteins.
 A/Reference number: S50754; MUID:95093034; PMID:8000007
 A/Accession: S50755
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-473 <MO>
 A/Cross-references: EMBL:L29029; NID:9530875; PIDN:AA53953.1; PID:9530876

Query Match
 Best Local Similarity 35.3%; Pred. No. 0.11; Length 473;
 Matchee 42; Conservative 12; Mismatches 51; Indels 14; Gaps 3;

Qy 88 RKGCKPLGELIAGTSSRYTPYSSQNSQPLSS-----AFQSPISYQVSSSSS 135
 Db 262 RTGASPSBSPKASBPVKVSPSPKASBPSPKASBPSPSPKASBPSPSPKASBPSP 321
 Qy 136 FPPSSRGPPNNMSTFFPLRNGGIPSLPLRLRISNCEPTVPVSSPT-SKNPKPLPN 193
 Db 322 SPSPS-PPKASPPSPSPSPVQSPKSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 379

RESULT 12

S46660

weta protein - Pentacillium chrysogenum
 C/Species: Pentacillium chrysogenum
 C/Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #ext_change 26-Aug-1999
 C/Accession: S46660
 R/Prade, R.A.; Timberlake, W.E. Mol. Gen. Genet. 244, 538-547, 1994
 A/Title: The Pentacillium chrysogenum and Aspergillus nidulans weta developmental regulat
 A/Reference number: S46660; MUID:94359480; PMID:8078481
 A/Accession: S46660
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-499 <PR>
 A/Cross-references: EMBL:X80058; NID:9510873; PIDN:CAA56364.1; PID:9510874
 C/Superfamily: regulatory protein weta

Query Match
 Best Local Similarity 7.5%; Score 132; DB 2; Length 499;
 Matches 91; Conservative 47; Mismatches 147; Indels 138; Gaps 19;

Qy 2 TSDGATSTSAAAAAAAAAARRK-----SWRENNRRRRRR----- 41
 Db 109 TASHGLST---PSTPATPRKKTQSLIRPKSIRHRSPERKSHLRKQSFPSLMRSS 165

Qy 42 --AVALKITGLAAGDYNLPKHCDDNNVILKALCVAAGVVEEDGTTTRKGCPLPGEI- 98
 Db 166 NLKAMAYPEMAAORLQNFSLH-----GSDRLPLSPSPDVL 204
 Qy 99 -----AGTSRRVTPYSSQ--NOSPSSAFOSPSPS-----YQVSP 131
 Db 205 ICHENMPTQIMNQHDSERASQYDARLYQSP---SVMPSPSLMARSROOQHYAOP 261
 Qy 132 SSSPSPSPRGEPNNMSTFFPLRNGGIPSLPLRLRISNCEPTVPVSSPTSK- 186
 Db 262 SSSSLTNSPSSADLFFSS-----HSSDPHLSWQSDPLHASSLFTPLDQGDQSQW 315
 Qy 187 -----NPKPLPMWESIAKQSMIAKQSMASFNYPFYAASA----- 221
 Db 316 WSPMPSRVAAQQGAYLTSPTVPTMQSVQSDMMQGLMIGN-PSYDMSADHSFSSN 374
 Qy 222 --PASPT-----HRHQFTLTATPECD-ESDSSTVDSGHWISFQFAQQPFSASWPT 272
 Db 375 MLPATPQKFTSFTNYSQVNVSRSPSLAPKAGTSPDRTRKWHFO--AHSSPHSQRL-S 431
 Qy 273 SPTFNLVLRPAQOMSPTAAFOEIGQSEFKFENSQVKPWEGERIDVGMEDLELTUNG 332
 Db 432 GGSMAAPKPAKAGSSSSRGSNSKSVSVS---FVNFTA-----HDSKILTELPSGSS 479
 Qy 333 KAR 335
 Db 480 KTK 482

RESULT 13

A27677

DNA-directed RNA polymerase (EC 2.7.7.6) II largest chain - Chinese hamster (fragment)
 C/Species: Crictetus griseus (Chinese hamster)
 C/Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #ext_change 18-Jun-1999
 C/Accession: A27677
 R/Allison, U.A.; Wong, J.K.C.; Fitzpatrick, V.D.; Moyle, M.; Ingles, C.J. Mol. Cell. Biol. 8, 321-329, 1988
 A/Title: The C-terminal domain of the largest subunit of RNA polymerase II of Saccharom.
 A/Reference number: A33104; MUID:88094402; PMID:3122024
 A/Accession: A27677
 A/Molecule type: DNA
 A/Residues: 1-467 <ALL>
 A/Cross-references: GB:M19538; NID:9191170; PIDN:AAA37008.1; PID:9387055
 C/Genetics:
 A/Gene: RPO21
 A/Introns: 33/1; 83/1
 C/Superfamily: human DNA-directed RNA polymerase II largest chain
 C/Keywords: DNA binding; nucleotidyltransferase; tandem repeat; transcription; zinc fin

Query Match
 Best Local Similarity 7.4%; Score 131; DB 2; Length 467;
 Matches 65; Conservative 23; Mismatches 91; Indels 40; Gaps 12;

Qy 93 PLPELIAGTSSRYTPYSSQNSQPLSSAFQSPISYQVSSSSSF-PPSPRGEPNNMST 151
 Db 261 PTSP 314
 Qy 152 FFPPLRNGGIPSLPLRLRISNCEPTVPVSSPTSKNPKP-----LPNWSIA-KQSM 202
 Db 315 YTP-----SSPSYSPS--SPSYSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 366
 Qy 203 AIAKQSMASFN-----FFYAVSADA-SPTNHHQFTLTATPECDSDSDSTVDS 250
 Db 367 TSPKSP 426
 Qy 251 GHWISFQFAQQPFSASWPTSPFNLVLRPAQOMSPN 289
 Db 427 -YSPTSPKSGSYSP 461

RESULT 14

A28490

DNA-directed RNA polymerase (EC 2.7.7.6) II largest chain - mouse

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2003, 14:24:08 / Search time 17 seconds
(without alignment)

929.469 Million cell updates/sec

Title: US-09-995-938A-7

Sequence: 1 MTSDGATSTSTMAAAAAAAAA.....IHDVGMEDLELTGNGKARG 336

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 127863 aeqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB aeq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|-----------------------|
| 1 | 146 | 8.3 | 1852 | 1 RPB1_CAEEL | P16356 caenorhabditis |
| 2 | 133.5 | 7.6 | 314 | 1 PIX2_BRARE | O9W522 brachydanio |
| 3 | 132 | 7.5 | 499 | 1 WETA_PENCH | O01870 penicillium |
| 4 | 131 | 7.4 | 467 | 1 RPB1_CRIGR | P11414 citreulius |
| 5 | 131 | 7.4 | 1970 | 1 RPB1_HUMAN | P24928 homo sapien |
| 6 | 131 | 7.3 | 1970 | 1 RPB1_MOUSE | P08775 mus musculu |
| 7 | 128.5 | 7.2 | 326 | 1 PIX2_XENLA | O9W713 xenopus lae |
| 8 | 127.5 | 7.2 | 1493 | 1 M3K1_RAT | O62925 rattus norv |
| 9 | 126.5 | 7.2 | 317 | 1 PIX2_HUMAN | O99697 homo sapien |
| 10 | 126.5 | 7.2 | 317 | 1 PIX2_MOUSE | P97474 mus musculu |
| 11 | 126.5 | 7.2 | 324 | 1 PIX2_RAT | O97074 rattus norv |
| 12 | 126.5 | 7.2 | 1733 | 1 RPB1_YEAST | P04050 saccharomyc |
| 13 | 125.5 | 7.1 | 1887 | 1 PIX2_CHICK | O93385 gallus gall |
| 14 | 124.5 | 7.1 | 1887 | 1 RPB1_DROME | P04052 drosophila |
| 15 | 124 | 7.0 | 1367 | 1 AMYH_YEAST | P08640 saccharomyc |
| 16 | 123.5 | 7.0 | 917 | 1 SMOC_HUMAN | P53814 homo sapien |
| 17 | 122.5 | 6.9 | 2414 | 1 P300_HUMAN | O09472 homo sapien |
| 18 | 121.5 | 6.9 | 855 | 1 GAF1_SCHPO | O10289 schizosacch |
| 19 | 121 | 6.9 | 806 | 1 MKO7_MOUSE | O9W948 mus musculu |
| 20 | 121 | 6.9 | 1752 | 1 RPB1_SCHPO | P18616 arabidopsis |
| 21 | 121 | 6.9 | 1840 | 1 RPB1_ARATH | P35084 dictyosteli |
| 22 | 120.5 | 6.8 | 438 | 1 RSV1_SCHPO | O9P768 schizosacch |
| 23 | 120.5 | 6.8 | 902 | 1 RPB1_DICDI | P53349 mus musculu |
| 24 | 120.5 | 6.8 | 1493 | 1 M3K1_MOUSE | P35084 dictyosteli |
| 25 | 120 | 6.8 | 1192 | 1 RTN4_HUMAN | O9W948 mus musculu |
| 26 | 118.5 | 6.7 | 845 | 1 KRUI_DROME | O9W948 mus musculu |
| 27 | 118 | 6.7 | 2090 | 1 N214_HUMAN | P35658 homo sapien |
| 28 | 118 | 6.7 | 2842 | 1 APC_FAT | P70478 rattus norv |
| 29 | 118 | 6.7 | 2845 | 1 APC_MOUSE | O61315 mus musculu |
| 30 | 117.5 | 6.7 | 292 | 1 PIX3_XENLA | O91883 xenopus lae |
| 31 | 117 | 6.6 | 1365 | 1 SUZ2_DROME | P55172 drosophila |
| 32 | 117 | 6.6 | 5038 | 1 PCLO_MOUSE | O9Q977 mus musculu |
| 33 | 116.5 | 6.6 | 565 | 1 SCRI_SCHPO | O14335 schizosacch |

ALIGNMENTS

RESULT 1
ID RPB1_CAEEL STANDARD: PRT, 1852 AA.
AC P16356; Q20090;
DT 01-AUG-1990 (Rel. 15, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA-directed RNA polymerase II large subunit (EC 2.7.7.6).
GN AMA-1 OR P3644.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=90066416; PubMed=2586513;
RT Bird D.M., Riddle D.L.;
RT "Molecular cloning and sequencing of ama-1, the gene encoding the
RT largest subunit of Caenorhabditis elegans RNA polymerase II.";
RL Mol. Cell. Biol. 9:4119-4130(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Miller N., Bradshaw H.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RA REVISIONS.
RA Waterston R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N-nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM NINE TO
CC FOURTEEN DIFFERENT POLYPEPTIDES.
CC -1- SUBCELLULAR LOCATION: Nucleus.
CC -1- PPM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
CC THE PHOSPHORYLATION ACTIVATES POL2.
CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND TRNA GENES.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>)
CC or send an email to license@ebi.ac.uk.
CC -----
DR EMBL; M29235; AAA28126.1; -;
DR EMBL; U53333; AAA96158.2; -;

DR PIR: A34092; A34092.
DR Nonmeped; F36A4.7; CE28300.
DR InterPro; IPR000722; RNA_pol_A.
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007060; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR007075; RNA_pol_Rpb1_6.
DR InterPro; IPR007073; RNA_pol_Rpb1_7.
DR InterPro; IPR006592; RNA_polA_N.
DR InterPro; IPR006684; RNA_polIT_repeat.
DR Pfam; PF04997; RNA_pol_Rpb1_1.
DR Pfam; PF006623; RNA_pol_Rpb1_2.
DR Pfam; PF04983; RNA_pol_Rpb1_3.
DR Pfam; PF05000; RNA_pol_Rpb1_4.
DR Pfam; PF04998; RNA_pol_Rpb1_5.
DR Pfam; PF04992; RNA_pol_Rpb1_6.
DR Pfam; PF04990; RNA_pol_Rpb1_7.
DR Pfam; PF05001; RNA_pol_Rpb1_R; 22.
DR SMART; SMO0663; RPOA_N.1.
DR PROSITE; PS00115; RNA_POL_11_REPEAT; 26.
DR Translatase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
KW DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
FT ZN FING 66 82
FT DOMAIN 1560 1852
FT CONFILCT 215 215
FT CONFILCT 911 911 R -> D (IN REF. 1).
FT CONFILCT 959 959 R -> RSVAKNAITL (IN REF. 1).
FT CONFILCT 974 974 Q -> L (IN REF. 1).
FT CONFILCT 990 991 KP -> NA (IN REF. 1).
FT CONFILCT 1156 1158 MISSING (IN REF. 1).
FT CONFILCT 1402 1403 IT -> YG (IN REF. 1).
SO SEQUENCE 1852 AA; 203978 MW; 21184563119088B CRC64;

Query Match 8.3%; Score 146; DB 1; Length 1852;
Best Local Similarity 27.5%; Pred. No. 0.05;
Matches 67; Conservative 34; Mismatches 95; Indels 48; Gaps 12;

| | | |
|----|--|--|
| RT | SEQUENCE FROM N.A. (ISOFORMS PITX2A AND PITX2C). | |
| RA | Yan Y.-T., Gitman K., Ding J., Burdine R.D., Corrales J., | |
| RA | Price S.M., Talbot W.S., Schier A.F., Shen M.M.; | |
| RT | "Conserved role for EGF-CFC genes in vertebrate left-right axis | |
| RT | formation."; | |
| RT | Submitted (JUN-1999) to the EMBL/GenBank/DDJ databases. | |
| RM | | |
| RP | SEQUENCE FROM N.A. (ISOFORMS PITX2A AND PITX2C). | |
| RA | Eisner J.J., Zhang J., Yost H.J.; | |
| RT | "Conserved isoforms of the pitx2 gene have differential activity and | |
| RT | expression in zebrafish." | |
| RL | Submitted (MAR-1999) to the EMBL/GenBank/DDJ databases. | |
| RM | | |
| RM | | |
| RP | SEQUENCE FROM N.A. (ISOFORM PITX2A). | |
| RA | MEDLINE=99432240; PubMed=10500184; | |
| RA | Tsukui T., Capdevila J., Tamura K., Ruiz-Lozano P., | |
| RA | Rodriguez-Esteban C., Yonet-Tamura S., Magallon J., | |
| RA | Chandrasekara R.A.S., Chien K., Blumberg B., Evans R.M., | |
| RA | Izpisua-Belmonte J.-C.; | |
| RT | "Multiple left-right asymmetry defects in Shh(-/-) mutant mice unveil | |
| RT | a convergence of the Shh and retinoic acid pathways in the control of | |
| RT | Lefty-1." | |
| RT | Proc. Natl. Acad. Sci. U.S.A. 96:11376-11381(1999). | |
| CC | - FUNCTION: May play an important role in development and | |
| CC | maintenance of anterior structures. | |
| CC | - SUBCELLULAR LOCATION: Nuclear. | |
| CC | - ALTERNATIVE PRODUCTS: | |
| CC | Event=Alternative splicing; Named isoforms=2; | |
| CC | Name=PITX2C; | |
| CC | Isoid=Q9W5Z2-1; Sequence=Displayed; | |
| CC | Name=PITX2A; | |
| CC | Isoid=Q9W5Z2-2; Sequence=VSP 002268, VSP 002269; | |
| CC | - SIMILARITY: BELONGS TO THE PAIRED HOMEBOX FAMILY. BICOID | |
| CC | SUBFAMILY. | |
| CC | - SIMILARITY: Contains 1 homeobox domain. | |
| CC | - SIMILARITY: Contains 1 OAR domain. | |
| CC | ----- | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | |
| CC | use by non-profit institutions as long as its content is in no way | |
| CC | modified and this statement is not removed. Usage by and for commercial | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | |
| CC | or send an email to license@isb-sib.ch). | |
| CC | ----- | |
| DR | EMBL; AF156905; AAD40179.1; - | |
| DR | EMBL; AF156906; AAD40180.1; - | |
| DR | EMBL; AF132446; AAD34390.1; - | |
| DR | EMBL; AF132447; AAD34391.1; - | |
| DR | EMBL; AF181681; AAF00486.2; - | |
| DR | HSSP; P06601; IFTL. | |
| DR | ZFIN; ZDB-GENE-990714-27; pitx2a. | |
| DR | InterPro; IPR003654; Homeo OAR. | |
| DR | InterPro; IPR001356; Homeobox. | |
| DR | InterPro; IPR007104; Paired_homeo. | |
| DR | Pfam; PF00046; homeobox_1. | |
| DR | Pfam; PF03826; OAR; 1. | |
| DR | PRINTS; PR00024; HOMEBOX. | |
| DR | ProDom; PD00010; Homeobox_1. | |
| DR | SMART; SM00389; HOX; 1. | |
| DR | PROSITE; PS00027; HOMEBOX_1; 1. | |
| DR | PROSITE; PS00071; HOMEBOX_2; 1. | |
| DR | PROSITE; PS50803; OAR; 1. | |
| KM | Homeobox; DNA-binding; Developmental protein; Nuclear protein; | |
| KM | Alternative splicing; | |
| FT | DNA BIND 82 141 | HOMEBOX. |
| FT | DOMAIN 276 289 | OAR. |
| FT | DOMAIN 282 286 | NUCLEAR LOCALIZATION SIGNAL (POTENTIAL). |
| FT | VARSPIC 1 45 | Misling (in isoform PITX2A). |
| FT | VARSPIC 46 60 | /FTId=VSP_002268. |
| FT | | LDVHVSPTSSPEV -> MSHCRKLASTCAQL (in |
| FT | | isoform PITX2A). |

FT SEQUENCE 314 AA; 35187 MW; 520DF7855664A6B5 CRC64; /PTID=VSP 002269.

Query Match 7.6%; Score 133.5; DB 1; Length 314;
 Best Local Similarity 26.3%; Pred. No. 0.049;
 Matches 61; Conservative 25; Mismatches 87; Indels 59; Gaps 11;

QY 6 ATSTSAATAAAAAARRKPSRENNRRRRRAVAAKIYTGIRAGQDYLPHGCON 65
 DB 116 AVMTLTERVAVWPKRRAKRRKNOAELCKGFGPO--FNGLMQPYDDMYPSITYN 174
 QY 66 NEVLKALCEAGWVEEDGTYRKCKPLPGELAGTSRVTY--YSSQNSPLSS--AFQS 122
 DB 175 NMAAKGL-----TSASLSTKSPFPFNSMNNPLSSQTMESP 210
 QY 123 PIPSVQVPPSSSPSPSGEPN-----NNMSSTFPPLRNGCIPSSLIRISNCPV 176
 DB 211 PMSISSMSMSSMVSATVGVPGSSILNLSNPL--NSGVPTP-----ACPY 261
 QY 177 TPVSSPTSKPKPLPNMESIAKQSMAL---AKQSMASFNYPFAVSPAS 224
 DB 262 APPT-----PPVYTRDTGNSLSLRLAKOH--SEFGVA--SVONPAS 301

RESULT 3
 WETA_PENCH STANDARD; PRT; 499 AA.

AC 001870;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Regulatory protein WETA.
 GN WETA.
 OS Penicillium chrysogenum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OX Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; penicillium.
 OX NCBI_TaxID=5076;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Maiconin 54-1255;
 RA MEDLINE=94359480; PubMed=8078481;
 RX Prade R.A., Timberlake W.E.;
 RT "The penicillium chrysogenum and Aspergillus nidulans weta
 RT developmental regulatory genes are functionally equivalent.";
 RL Mol. Genet. 244:539-547(1994).
 CC -1- FUNCTION: RESPONSIBLE FOR ACTIVATING A SET OF GENES WHOSE PRODUCTS
 CC MAKE UP THE FINAL TWO CONIDIAL WALL LAYERS OR DIRECT THEIR
 CC ASSEMBLY AND THROUGH THIS ACTIVITY IS RESPONSIBLE FOR ACQUISITION
 CC OF SPORE DOMINANCY. REGULATES THE EXPRESSION OF SPORE-SPECIFIC
 CC GENES.
 CC -1- DOMAIN: HAS AN ACIDIC N-TERMINUS (AA 1-50) FOLLOWED BY A SER-,
 CC THR-, PRO-RICH DOMAIN (AA 118-218) AND A BASIC C-TERMINUS (AA
 CC 428-499).
 CC -1- SIMILARITY: TO E.NIDULANS WETA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.1db-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; X80058; CA56364.1; -;
 DR PIR; S46660; S46660.
 KW Developmental protein; Conidiation; Transcription regulation;
 FT Activator.
 FT DOMAIN 262 265 POLY-SER.
 FT SEQUENCE 499 AA; 54899 MW; 0CC66F4E15248107 CRC64;

Query Match 7.5%; Score 132; DB 1; Length 499;
 Best Local Similarity 21.5%; Pred. No. 0.1;
 Matches 91; Conservative 47; Mismatches 147; Indels 138; Gaps 19;

QY 2 TSDGATSTSAATAAAAAARRKP-----SWRENNRRRRRRR----- 41
 DB 109 TASHGLSTG-----PSTPPATPRRRKPTQSALITPRKSIHRSFNEARSHLRKQFSPSLMRSS 165
 QY 42 --AVAKIYTGIRAGQDYLPHGCONNEVLKALCEAGWVEEDGTYRKCKPLPGELI- 98
 DB 166 NUSKARMAYPEAMARLQNFSLH-----GSEDRLLPSPPSDVL 204
 QY 99 -----AGTSRVTPYSSQ--NOSPASAFQSPIS-----QVSP 131
 DB 205 IOHEMPTBQINQMGDSERASQIDARLYQSP--SYMSPSIAMSARQQOHYAOP 261
 QY 132 SSSSFSPSPSGEPNNMSSTFPPLRNGCIPSPSLRLI-----SNSCPVTPVSSPTSK- 186
 DB 262 SSSSLTNSSPSADDFSS-----HSSDPHLSLWQSDPLHSLSLSTPDLQGDQSQW 315
 QY 187 -----NPKPLPNMESIAKQSMALAKQSMASFNYPFAVSA----- 221
 DB 316 WSPMSRYAQQGAYLTPTPTVTRWQSVQSDMMQGLMIQFN--PSYMSADHSFSSN 374
 QY 222 --PASPT-----HRHOFHLATIPEDC--ESDSTVDSGHWISFQKFAQQQPPSASMPV 272
 DB 375 MLPATPOKFDTSFNTSQVHNVSRSPSLSPKATSPDTRKWHFO--AHSSPHSQRL-S 431
 QY 273 SPTFNLVAPAPQOMSPTAAFOELIGQSSSPKFNQYKWEGERIHDVGMEDLELTGNG 332
 DB 432 GQSMNAPKPAKASGSSSKGSKSVS-----FVNFTA-----HDSKILTELPSSGS 479
 QY 333 KAR 335
 DB 480 KTK 482

RESULT 4
 RPB1_CRIGR STANDARD; PRT; 467 AA.

AC 011414;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6) (RPB1
 DE (fragment).
 GN POLR2A.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetulus.
 OX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=86094402; PubMed=3122024;
 RA Allison L.A., Wong J.K.-C., Fitzpatrick V.D., Moyls M., Ingles C.J.;
 RT "The C-terminal domain of the largest subunit of RNA polymerase II of
 RT Saccharomyces cerevisiae, Drosophila melanogaster, and mammals: a
 RT conserved structure with an essential function.";
 RL Mol. Cell. Biol. 8:321-329(1988).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC -1- SUBUNIT: RNA POLYMERASE II CONSISTS OF 12 DIFFERENT SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
 CC THE PHOSPHORYLATION ACTIVATES POL2.
 CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
 CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
 CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
 CC III FOR 5S AND TRNA GENES.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as the content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).

CC -----

DR EMBL, M19538; AAA37008.1; -

DR PIR, A27677; A27677.

DR InterPro; IPR00684; RNA_polII_repeat.

DR Pfam; PF05001; RNA_pol_Rp1i_R; 28.

DR PROSITE; PS00115; RNA_POL_II_REPEAT; 43.

KM Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;

CC DNA-binding; Nuclear protein; Phosphorylation.

FT NON TER

FT DOMAIN

FT SEQUENCE 467 AA; 48167 MW; 0AADD923B5D2BC CAC64;

Query Match 7.4%; Score 131; DB 1; Length 467;

Best Local Similarity 29.7%; Pred. No. 0.11;

Matches 65; Conservative 23; Mismatches 91; Indels 40; Gaps 12;

QY 93 PLPEGIACTSKRVTPYSSQONOSPLSSAPQSPISQVSPSSGSF-PSPSRGEPNNMST 151

Db 261 PTPSPSPSTPSPVYTP-TSPNVSPSTSPSPSPSPSY--SPTSPSPSPSPRPYTP--QSP 314

QY 152 PPPPLRNGCIPSSLSLRISNGCPPTPPVSPSTSKNRP-----LPNMESIA-KQSM 202

Db 315 YTP-----SP 366

QY 203 AIAKQSMASFNV-----PPYAVASAP-SPTHRHQFTLATIPEDCESDSTYDS 250

Db 367 TSPKISPSPTSP 426

QY 251 GHWISFOKFAQQPSPASNVPTSPFTNLVKPAPQOMSPN 289

Db 427 -YSPTPSPKQSTVSPSTSP 461

RESULT 5

RPBI_HUMAN

ID RPBI_HUMAN STANDARD; PRT; 1970 AA.

AC P24928;

DT 01-MAR-1992 (rel. 21. Created)

DT 01-MAR-1992 (rel. 21. Last sequence update)

DT 16-OCT-2001 (rel. 40. Last annotation update)

DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6) (RPBI).

GN POLR2A.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=92178992; PubMed=1542581;

RA Wintzerich M., Acker J., Vicaire S., Vigneron M., Kedinger C.;

RT "Complete sequence of the human RNA polymerase II largest subunit";

RL Nucleic Acids Res. 20:910-910(1992).

RN [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=95347616; PubMed=7622068;

RA Mita K., Tsuji H., Moriyomo M., Takahashi E., Nenoi M.,

RA Ichimura S., Yamachi M., Hongo E., Hayashi A.;

RT "The human gene encoding the largest subunit of RNA polymerase II.";

RL Gene 159:285-286(1995).

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION

CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS

CC SUBSTRATES.

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

CC {RNA} (N).

CC -1- SUBUNIT: RNA POLYMERASE II CONSISTS OF 12 DIFFERENT SUBUNITS.

CC -1- SUBCELLULAR LOCATION: Nucleus

CC -1- PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.

[illegible]

```

0y 251 GHWISFQKFAQQCFSA5MPT5TSTFNLVVPAPQSPEN 289
Db 1930 -YSPSPKSGT5PTSPG5PTSP5T5TSPALTSIPA--ISP 1964

RESULT 6
RPB1_MOUSE ID STANDARD; PRT; 1970 AA.
AC P08775;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6) (RPB1).
OS POLR2A OR RPO2-1 OR RPI1215.
OS Mus musculus (Mouse).
OC Burkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE=87280135; PubMed=3038894;
RX Ahearn J.M. Jr., Bartolomei M.S., West M.L., Cisek L.J., Corden J.L.;
RT "Cloning and sequence analysis of the mouse genomic locus encoding
RL the largest subunit of RNA polymerase II.";
RL J. Biol. Chem. 262:10695-10705(1987).
RP (2)
RP SEQUENCE OF 1587-1970 FROM N.A.
RX MEDLINE=86068017; PubMed=2999785;
RA Corden J.L., Cadena D.L., Ahearn J.M. Jr., Dahmus M.E.;
RT "A unique structure at the carboxyl terminus of the largest subunit
RL of eukaryotic RNA polymerase II.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7934-7938(1985).
RP (3)
RP REVISIONS, AND PRESENCE OF AN ADDITIONAL EXON.
RX MEDLINE=92178992; PubMed=1542581;
RA Wintzerich M., Acker J., Vicaite S., Vigneron M., Kedinger C.;
RT "Complete sequence of the human RNA polymerase II largest subunit.";
RL Nucleic Acids Res. 20:910-910(1992).
RP (4)
RP FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
OF DNA INTO RNA USING THE FOUR RIBONUCLOSIDE TRIPHOSPHATES AS
SUBSTRATES.
RP (5)
RP -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA] (N).
RP (6)
RP -1- SUBUNIT: RNA POLYMERASE II CONSISTS OF 12 DIFFERENT SUBUNITS.
RP (7)
RP -1- SUBCELLULAR LOCATION: Nuclear.
RP (8)
RP -1- PIV: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
RP (9)
RP THE PHOSPHORYLATION ACTIVATES POL2.
RP (10)
RP -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
III FOR 5S AND TRNA GENES.
RP (11)
RP -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
RP (12)
RP THIS SWISS-PROT entry is a copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
RP (13)
RP -----
DR EMBL: M12130; AAA40071.1;
DR EMBL: M14101; AAA40071.1; JOINED.
DR PIR: A28490; A28490.
DR MCD: MGI:98086; Polr2a.
DR GO: GO:0005634; C:nucleus; ISS.
DR GO: GO:0003677; F:DNA binding activity; ISS.
DR GO: GO:0003901; F:DNA-directed RNA polymerase II activity; ISS.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; ISS.
DR GO: GO:0006365; P:transcription from Pol II promoter; ISS.
DR INRATPRO: IPR000722; RNA_pol_A.
DR InterPro: IPR007080; RNA_pol_Rpb1_1.

```

```
DR InterPro: IPR007066; RNA_pol_Rpb1_3
DR InterPro: IPR007083; RNA_pol_Rpb1_4
DR InterPro: IPR007081; RNA_pol_Rpb1_5
DR InterPro: IPR007075; RNA_pol_Rpb1_6
DR InterPro: IPR007073; RNA_pol_Rpb1_7
DR InterPro: IPR006592; RNA_pola_N.
DR InterPro: IPR006684; RNA_poliI_repeat.
DR Pfam: PF04997; RNA_pol_Rpb1_1; 1.
DR Pfam: PF00623; RNA_pol_Rpb1_2; 1.
DR Pfam: PF04983; RNA_pol_Rpb1_3; 1.
DR Pfam: PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam: PF04998; RNA_pol_Rpb1_5; 1.
DR Pfam: PF04992; RNA_pol_Rpb1_6; 1.
DR Pfam: PF04990; RNA_pol_Rpb1_7; 1.
DR Pfam: PF05001; RNA_pol_Rpb1_R; 29.
DR SMART; SMO0663; RPOLA N. 1.
DR PROSITE; PS00115; RNA_POL II REPEAT; 42.
KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
KM DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
FT ZN FING 71 87 C2H2-TYPE (POTENTIAL).
FT DOMAIN 1590 1958 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
FT CONFLICT 1498 1498 P -> R (IN REF. 1 AND 2).
FT CONFLICT 1499 1536 MISSING (IN REF. 1 AND 2).
SQ SEQUENCE 1970 AA; 217175 MW; 7D6F3BF9D2A657E CR664;

.Query Match
Best Local Similarity 7.4%; Score 131; DB 1; Length 1970;
Matches 65; Conservative 23; Mismatches 91; Indels 40; Gaps 12

OY 93 PLPEGIAGTSRVTVPVSSQNSQLSSAFQSIPSYCYSPSSSR-PESPGRKPPNNMSST 151
DB 1764 PTSGSYSPTSNTYP-TSPNVSPSPSPSPSPSPSY--SPTPSPSPSPSPRYTP--OSPT 1817
OY 152 FFPLRLNGCIGPS$LPISLRINSCEPVSPVSPTSINKPKP-----LPNWESIA-QSM 202
DB 1818 YTF-----SESYSYPS---SPSYSPITPKXTPTSP$YSPPSPPEYTPASPKXSPSXSP 1866
OY 203 AIAKQSAASFNY-----PFVAVASAPA-SPTHROPHLTATIECEDSDSVDS 250
DB 1870 TSPKYSPSTPYPTTKPKYSPTPSYSPVTVTSPKXSPSPSPSPSPSPSPSPSPSPSP 1929
OY 251 GHMISFOKAQQCPFSASMYTSTETFLVKPADQONSPN 289
DB 1930 -YSTSPKGGTYGPTSGYSPSTPYSLTSPA---TSPD 1964

RESULT 7
PIXZ_XENLA STANDARD; PRT; 326 AA.
AC Q9PMW3; O9J396; Q9YHA4; Created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DE 28-FEB-2003 (Rel. 41, last annotation update)
GN PIXZ OR PIX2.
OS Xenopus laevis (African clawed frog).
OC Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Eumetastomi;
CX xenopodinae; xenopus.
NCBI_TaxID=8355.
[1]
SEQUENCE FROM N.A. (ISOFORM PIX2A).
MEDLINE=98370612; PubMed=9707115;
RA Ryan A.K., Blumberg B., Rodriguez-Berehan C., Yonel-Tanura S.,
RA Tamura K., Taubki T., de la Pena J., Sabbagh W., Greenwald J.,
RA Choe S., Norris D.P., Robertson E.O., Evans R.M., Rosenfeld M.G.,
RA Tzipisua-Belmonte J.-C.;
RT Pixz determines left-right asymmetry of internal organs in
vertebrates."
RL Nature 394:545-551(1998).
```


DR PROSITE, P550089; ZF-RING 2; 1.
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KM Phosphorylation; Zinc-finger.
 FT ZN PING 433 482
 FT DOMAIN 1224 1485 RING-TYPE
 FT DOMAIN 25 29 PROTEIN KINASE
 FT DOMAIN 74 149 POLY-GLY
 FT DOMAIN 233 291 PRO-RICH
 FT DOMAIN 412 421 PRO-RICH
 FT DOMAIN 1163 1168 POLY-SER
 FT NP BIND 1230 1238 POLY-GLU
 FT BINDING 1253 1253 ATP (BY SIMILARITY)
 FT ACT_SITE 1350 1350 ATP (BY SIMILARITY)
 FT MUTAGEN 1369 1369 BY SIMILARITY
 SC SEQUENCE 1493 AA; 161315 MW; 8C5F29F6668524 CRC64;
 Query Match 7.2%; Score 127.5; DB 1; Length 1493;
 Bees Local Similarity 22.2%; Pred. No. 0.68;
 Matches 76; Conservative 38; Mismatches 129; Indels 99; Gaps 12;
 QY 5 GATSTSAATAAAAAAARAKPSWRENNRRRRRAVAALITYGLRAQGYNLPMKMD 64
 DB 134 GATSPAGABPPPAAL-----PSGRENNK-----ETLKLHMKMD--RPERRM 174
 QY 65 NNEVLKALCVKGVVEEDGTTRKG---CKPLPGIAGTSSRVTPYSSQNSPLSSAFQ 121
 DB 175 IRELUKATCMRA-WGHEMLERNRNRKGPVVKPIP--IKGDSGSMNLAEIQEGQAGSA 231
 QY 122 SPIBQVQVSSSSSPSSRGRBNNNMSSTFPPPLRNGIPSLPLSRISNCPV----- 176
 DB 232 AAPAGRRSPSGSSPSGRSGKPS-----FGVRRKRVSPVPFQSG 272
 QY 177 --TPVSSPTSKNPL-PMNESIKQSMALQSMASNYFVAVSAPSTR----- 228
 DB 273 RITPRRABPDGFSYPSPETSRVNVKVMARLYLQIQIGNSFLIGDSDPNKRVPI 332
 QY 229 -----HQFTLATIPEDSD-----SSTYDS 250
 DB 333 GPQNSCGRGFTCHLVMLRVFPLEPSDPLMKTKLKNFEVESLFOKYSRRSRRIKA 392
 QY 251 GHWISFQKFAOQ---QPPSASMVPTSPPLNLPKAPQOMSP 288
 DB 393 PSRNTIQKFVSRMSNCHTLSSSTSTSSSENSIKQEBQMC 434
 RESULT 9
 PIX2 HUMAN STANDARD: PRT. 317 AA.
 AC 099697; 060578; 060579; 060580; Q9BY17;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Plutary homeobox 2 (RIGB bicoid-related homeobox transcription factor) (Soluble) (ALL1 responsive protein ARP1).
 GN PITX2 OR RIEG1 OR RIEG OR RGS OR ARP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND VARIANTS RIEG1 GLN-100; PRO-114 AND PRO-137.
 RC TISSUE=Craniofacial, and fetal brain;
 RA MEDLINE=97099449; PubMed=8944018;
 RA Semina E.V., Reiter R., Leyens N.J., Alward W.L.M., Small K.W.,
 RA Datsen N.A., Siegle-Bartelt J., Bleske-Nelson D., Bitoun P.,
 RA Zabel B.U., Carey J.C., Murray J.C.,
 RT Cloning and characterization of a novel bicoid-related homeobox
 RT transcription factor gene, RIEG, involved in Rieger syndrome.";
 Nac. Genet. 14:392-399(1996).
 RN [2]
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RA MEDLINE=98208582; PubMed=9539779;
 RA Arakawa H., Nakamura T., Zhadanov A.B., Fidanza Y., Yano T.,

RA Bullrich F., Shimizu W., Blechman J., Mazo A., Canaan E.,
 RA Croce C.M.;
 RT "Identification and characterization of the ARP1 gene, a target for
 RT the human acute leukemia ALL1 gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:4573-4578(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM PITX2A).
 RA Semina E.V., Funkhauser C., Bitoun P., Daack-Hirsch S., Alward W.L.M.,
 RA Amend B., Murray J.C.;
 RT "Spectrum and frequency of PITX2 mutations in patients with Rieger
 RT syndrome and related ocular anomalies.";
 RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM PITX2C).
 RC TISSUE=Lung;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buelow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stadelon M., Soares M.B., Bonaldo M.P., Caavaant I.L., Schetz T.E.,
 RA Bronsrein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywicki M.I., Skalska U., Smalls D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP VARIANT IRID2 TRP-89.
 RX MEDLINE=98099967; PubMed=9437321;
 RA Alward W.L.M., Semina E.V., Kelenak J.W., Heon E., Sheth B.P.,
 RA Stone E.W., Murray J.C.;
 RT "Autosomal dominant iris hypoplasia is caused by a mutation in the
 RT Rieger syndrome (RIGB/PITX2) gene.";
 RL Am. J. Ophthalmol. 125:98-100(1998).
 RN [6]
 RP VARIANT IRID2 HIS-115.
 RX MEDLINE=98282237; PubMed=9618168;
 RA Kulak S.C., Kozloweki K., Semina E.V., Pearce W.G., Walter M.A.;
 RT "Mutation in the RIEG1 gene in patients with iridogoniodysgenesis
 RT syndrome.";
 RL Hum. Mol. Genet. 7:1113-1117(1998).
 CC -1- FUNCTION: May play an important role in development and
 CC maintenance of anterior structures. Isoform PITX2C is involved in
 CC left-right asymmetry. The developing embryo (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=PITX2B; Synonyms=ARP1B;
 CC IsoId=Q99697-1; Sequence=Displayed;
 CC Name=PITX2C; Synonyms=ARP1C;
 CC IsoId=Q99697-2; Sequence=VSP_002260;
 CC Name=PITX2A; Synonyms=ARP1A;
 CC IsoId=Q99697-3; Sequence=VSP_002261;
 CC -1- DISEASE: Defects in PITX2 are the cause of Rieger syndrome type 1
 CC (RIGB) (MIM:180500). It is an autosomal dominant defect
 CC characterized by hypodontia (partial anodontia), anal stenosis,
 CC hypertelorism, mental deficiency, agenesis of the facial bones,
 CC with malformation of the anterior chamber of the eye.
 CC -1- DISEASE: Defects in PITX2 are the cause of iridogoniodysgenesis
 CC type 2 (IRID2) (MIM:137600); also known as iridogoniodysgenesis
 CC syndrome 2 (IGDS2). It is an autosomal inherited disease.
 CC -1- DISEASE: Defects in PITX2 are a cause of Peters anomaly
 CC (MIM:604229); a congenital defect of the anterior chamber of the

[illegible]

```

Db 119 AAVTNLLEARFVWFKRRAAKMRGRNRQALCKNGFPG-Q_FNLGMQYDDMYPGYSYN 177
QY .66 NEVIALCVAEGKAVVEDGDTYTRKGCCKPLPGEIACTSRRVP-YSSQNSPLS--SAFOS 122
Db 178 NMAAKGL-----TSASLSTKSPFFFNMMNVPPLSSQSMFSP 213
QY 123 PIPSYQVSPSSSPSPSRGEPN-----NNMSSTFFPFLRNGGSPSLPSLRISNCPV 176
Db 214 PMSISSMSMSSSWVPASVATGVFGSSLNLNMLNLSSTSL-NSAVPTP-----ACPV 264
QY 177 TPVVSPTSKNPKPLPNMESIAKOSMAI-----AKOSMASFNPYFVAVSAPAS 224
Db 265 APPI-----PPYVRDTCNSSLALRLKAKQH-SSEGYA--SVQNPAS 304

RESULT 10
P1X2_MOUSE STANDARD: PRT; 317 AA.
ID P1X2_MOUSE ST00336; P97933; Q9JUL0; Q9QXB8; Q9RLV9; Q9Z141;
AC P97474; O08646; O70336; P97933; Q9JUL0; Q9QXB8; Q9RLV9; Q9Z141;
DT 15-JUL-1998 (Rel. 36, created)
DT 15-JUL-1999 (Rel. 38, last sequence update)
DT 15-SEP-2003 (Rel. 42, last annotation update)
DE P1X2 homeobox 2 (orthodenticle-like homeobox 2) (Sourashin) (ALL1
DE responsive protein ARPI) (BRX1 homeoprotein) (Paired-like homeodomain
DE transcription factor Munc 30).
DE P1X2 OR P1X2 OR OTLX2 OR RGS OR ARPI OR BRX1.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN RP SEQUENCE FROM N.A. (ISOFORM PTX2A).
RA Muchicelli M.L., Martinez S., Patcyun A., Gontidis C., Brunet J.F.;
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
RN RN SEQUENCE FROM N.A. (ISOFORMS PTX2A AND PTX2B).
RC TISSUE-Pituitary;
RX MEDLINE=97227292; Pubmed=9147650;
RA Gage F.J., Camper S.A.;
RT Pituitary homeobox 2, a novel member of the bicoid-related family of
RT homeobox genes, is a potential regulator of anterior structure
RT formation";
RL Hum. Mol. Genet. 6:457-464(1997).
RN RN SEQUENCE FROM N.A. (ISOFORMS PTX2B AND PTX2C).
RP MEDLINE=98006140; Pubmed=9347917;
RA Kikimura K., Miura H., Yanazawa M., Miyashita T., Kato K.;
RT "Expression patterns of Brx1 (Rieg gene), Sonic hedgehog, Nkx2.2, Dlx1
RT and Arx during zona limitans intrathalamica and embryonic ventral
RL lateral geniculate nucleus formation.";
RL Mech. Dev. 67:83-96(1997).
RN RN SEQUENCE FROM N.A. (ISOFORMS PTX2A AND PTX2C).
RP MEDLINE=98208582; Pubmed=9539779;
RA Arakawa H., Nakamura T., Zhechanov A.B., Ridanza Y., Yano T.,
RA Bullrich F., Shimizu M., Blechman J., Mazo A., Canaan E.,
RA Croce C.M.;
RT "Identification and characterization of the ARPI gene, a target for
RT the human acute leukemia ALL1 gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4573-4578(1998).
RN RN PARTIAL SEQUENCE FROM N.A. (ISOFORM PTX2C).
RX MEDLINE=20054091; Pubmed=10585561;
RA Schwiebert A., Campbell M., Steinbeisser H., Blum M.;
RT "Pitx2 isoforms: involvement of Pitx2c but not Pitx2a or Pitx2b in
RL vertebrate left-right asymmetry.";
RL Mech. Dev. 90:41-51(2000).
RN RN PARTIAL SEQUENCE FROM N.A. (ISOFORM PTX2C).
RX MEDLINE=21220725; Pubmed=11319841;
RA Nicholson L.F.B., Mall, Gouding M.;
RT "Cloning and expression of Munc 30: a member of the paired-like
RL homeodomain gene family.";

```


DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6) (B2420).
 OS Eukaryotes; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 RX PubMed:8972577; Pubmed=8972577;
 RA Allison L.A., Moyle M., Shales M., Ingles C.J.;
 RT "Extensive homology among the largest subunits of eukaryotic and
 RT prokaryotic RNA polymerases."; Cell 42:599-610(1985).
 RL Cell 42:599-610(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX Medline=97127826; Pubmed=8972577;
 RA Woelfl S., Haneman V., Saluz H.B.;
 RT "Analysis of a 26,756 bp segment from the left arm of yeast
 RT chromosome IV."; Yeast 12:1549-1554(1996).
 RL Yeast 12:1549-1554(1996).
 RN [3]
 RP SEQUENCE OF 1669-1733 FROM N.A.
 RC STRAIN=S288C;
 RX Medline=95377607; Pubmed=7649444;
 RA Cronan J.E. Jr., Wallace J.C.;
 RT "The gene encoding the biotin-apolipoprotein ligase of Saccharomyces
 RT cerevisiae.";
 RL FEBS Microbiol. Lett. 130:221-230(1995).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC -1- SUBUNIT: RNA POLYMERASE II CONSISTS OF 12 DIFFERENT SUBUNITS.
 CC THIS SUBUNIT IS THE LARGEST COMPONENT OF RNA POLYMERASE II.
 CC -1- SUBCELLULAR LOCATION: Nucleus.
 CC -1- PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
 CC THE PHOSPHORYLATION ACTIVATES POL2.
 CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
 CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
 CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
 CC III FOR 5S AND 5.8S GENES.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on items
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-ibb.ch).
 CC -----
 DR EMBL: X03128; CAA26904.1; -
 DR EMBL: X96876; CAA65619.1; -
 DR EMBL: 274188; CAA98713.1; -
 DR EMBL: U27182; AAC49058.1; -
 DR PIR: S67866; RNBY2L.
 DR PDB: 1J30; 16-JUL-01.
 DR PDB: 1J50; 13-JUN-01.
 DR PDB: 1K83; 22-MAY-02.
 DR SGD: S0002299; RPO21.
 DR InterPro: IPR000722; RNA_pol_A.
 DR InterPro: IPR007080; RNA_pol_Rpb1_1.
 DR InterPro: IPR007065; RNA_pol_Rpb1_3.
 DR InterPro: IPR007083; RNA_pol_Rpb1_4.
 DR InterPro: IPR007081; RNA_pol_Rpb1_5.
 DR InterPro: IPR007073; RNA_pol_Rpb1_6.
 DR InterPro: IPR007075; RNA_pol_Rpb1_7.
 DR InterPro: IPR006592; RNA_polA_N.
 DR InterPro: IPR006584; RNA_polA_repeat.
 DR Pfam: PF04997; RNA_pol_Rpb1_1; 1.

DR Pfam; PP00623; RNA_pol_Rpb1_2; 1.
 DR Pfam; PP04983; RNA_pol_Rpb1_3; 1.
 DR Pfam; PP05000; RNA_pol_Rpb1_4; 1.
 DR Pfam; PP04998; RNA_pol_Rpb1_5; 1.
 DR Pfam; PP04992; RNA_pol_Rpb1_6; 1.
 DR Pfam; PP04990; RNA_pol_Rpb1_7; 1.
 DR Pfam; PP05001; RNA_pol_Rpb1_R; 14.
 DR SMART; SMO0663; RPOA_N; 1.
 DR PROSITE; PS00115; RNA_POL_II_REPEAT; 22.
 DR Transferrase; DNA-directed RNA polymerase; Zinc; Repeat;
 KM DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger;
 KM 3D-structure.
 FT ZN-FING 67 83 C2H2-TYPE (POTENTIAL).
 FT DOMAIN 1544 1719 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
 FT VARIANT 1553 1659 MISSING (IN STRAIN A364A).
 FT CONFLICT 1514 1514 A -> V (IN REF. 1).
 FT CONFLICT 1524 1524 G -> A (IN REF. 1).
 FT CONFLICT 1601 1601 T -> M (IN REF. 1).
 SQ SEQUENCE 1733 AA; 191610 MW; A45C1360FP9968 CRC64;
 Query Match 7 2%; Score 126.5; DB 1; Length 1733;
 Best Local Similarity 28.1%; Pred. No. 0.94;
 Matchea 64; Conservative 22; Mismatch 91; Indels 51; Gaps 10;
 QY 76 AGWVVEBDGTYRKCKPL-----PGEIAGTSRTVPSQNSQPLSAFQSPIS 126
 DB 1505 AGGTATGAGAYGERTSPFAGAYGAPTSG--FGVSSFGFSPTSTPTSTPSTPS 1562
 QY 127 YQVSPSSSSF-----PSPSRGEENNMSTFPPLNNGIPSLPLRLSNQCPVTPVSS 182
 DB 1563 Y--SPTSPSPSPSTPSPTSPSTPSPTSPSTPSPTSPSTPSPTSPSTPS 1606
 QY 183 PTKPKPKPLPMWESIAKOSMAIAKOSMASFNYFPVAVSAPA-SFPHRIQFHTLATIPCD 241
 DB 1607 PTS--PSYSPSPSPSPSTPSPTSPSTPSPTSPSTPSPTSPSPSPSTPSPTSPSTPS 1658
 QY 242 ESDSSTVDSGHMISFQKFAQQQPSASMVPTSPFTNLVKPAQOMSPN 289
 DB 1659 PAYSPTSPS-----YSPTSPSPSPSTPSPTSPSPSTPSPTSPN 1695
 RESULT 13
 ID PIX2_CHICK STANDARD; PRT; 333 AA.
 AC 093385; 073854; 09PWE0;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ptitutary homeobox 2 (PITX2).
 GN PITX2 OR PITX2.
 OS Gallus gallus (Chicken).
 OC Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 RX MEDLINE=98300291; Pubmed=9636662;
 RC STRAIN=White leghorn;
 RA St. Amant T.R., Ra J., Zhang Y., Hu Y., Baber S.I., Qiu M., Chen Y.;
 RA "Cloning and expression pattern of chicken Pitx2: a new component in
 the SHH signaling pathway controlling embryonic heart looping.";
 RA Biochem. Biophys. Res. Commun. 247:100-105(1998).
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Ryan A.K., Blumberg B., Rodriguez-Esteban C., Yonei-Tamura S.,
 RA Tamura K., Tsukitani T., de la Pena J., Sadegh W., Greenwald J.,
 RA Choe S., Norris D.P., Robertson E.J., Evans R.M., Rosenfeld M.G.,
 RA Ispilava-Belmonte J.C.;
 RA "Pitx2 determines left-right asymmetries in vertebrates.";
 RA Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RN SEQUENCE OF 170-319 FROM N.A.

RA Albajar M., Piedra M.E., Icardo J.M., Roe M.A., Rodriguez-Rey J.C.;
 RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: May play an important role in development and
 CC maintenance of anterior structures. May play a role in determining
 CC left-right asymmetry and in vasculogenesis during avian
 CC embryogenesis.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: Nuclear.
 CC Event=Alternative splicing; Named isoform=2;
 CC Name=1;
 CC IsoId=093385-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=093385-2; Sequence=VSP_002266; VSP_002267;
 CC -1- SIMILARITY: BELONGS TO THE PAIRED HOMEBOX FAMILY. BICOID
 CC SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 homeobox domain.
 CC -1- SIMILARITY: Contains 1 OAR domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF076640; AAC27322.1; -
 CC EMBL; AF077092; AAD46097.1; -
 CC EMBL; AF063935; AAC18346.1; -
 CC PIR; J00097; J00097.
 CC HSP; P06601; IYFL.
 CC InterPro; IPR003654; Homeo_OAR.
 CC InterPro; IPR001356; Homeobox.
 CC InterPro; IPR007104; Paired_homeo.
 CC Pfam; PF00046; homeobox; 1.
 CC Pfam; PF03826; OAR; 1.
 CC PRINTS; PR00024; HOMEBOX.
 CC ProDom; PD000010; Homeobox; 1.
 CC SMART; SMO0389; HOX; 1.
 CC PROSITE; PS00027; HOMEBOX_1; 1.
 CC PROSITE; PS00071; HOMEBOX_2; 1.
 CC PROSITE; PS00803; OAR; 1.
 CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 CC Alternative splicing.
 CC FT DNA_BIND 101 160 HOMEBOX.
 CC FT DOMAIN 101 160 OAR.
 CC FT DOMAIN 301 305 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DOMAIN 22 33 POLY-SER.
 CC FT VARSPLIC 1 62 Missing (in isoform 2).
 CC FT VARSPLIC 63 77 LEVHTISDTSPEEA -> MESNCRKLVASCVQL (in
 CC isoform 2).
 CC FT VARSPLIC 63 77 /FtId=VSP_002266.
 CC FT CONFLICT 234 234 S -> T (IN REF. 3).
 CC FT CONFLICT 292 292 T -> M (IN REF. 3).
 CC FT CONFLICT 317 317 M -> K (IN REF. 3).
 CC FT CONFLICT 328 328 P -> A (IN REF. 2).
 SQ SEQUENCE 333 AA; 36557 MW; FE9C68E475F5C85 CRC64;
 Query Match 7 1%; Score 125.5; DB 1; Length 333;
 Best Local Similarity 25.9%; Pred. No. 0.17;
 Matches 60; Conservative 27; Mismatch 86; Indels 59; Gaps 11;
 QY 6 ATSTSAATAAATAAARRKPSRENNRRRRRAVAAYKIYTLGACQVNLKHCQN 65
 DB 135 AVMTNLTEARVRVFKNRRAMKRRNQALCNGGCPQ-FNGLWQPYDMMYGYSTN 193
 QY 66 NEVLKALCVAGWVVEBDGTYRKCKPLPEIAGTSRTVTP-YSSQNSQPLS--SAFOS 122
 DB 194 NMAKGL-----TSASLSTKSPFPFNSMNVNPLSSQSMSP 229
 QY 123 PISQVSPSSSSSPSPSRGEPN-----NMSTTFPFLNNGIPSLPLRLSNQCPV 176

DB 230 ENSISMSMSMSVBPATVGVPCSSGLNSLNLN-----LSNPSLNSAVP---TPACPY 280

QY 177 TTPVSSPTSKNPKPLPNWESIAKQSMAI-----AKQSMASFPYVAVSAPAS 224

DB 281 AAPT-----PPYVRYDTGNSLSLAKKAKQH--SSFGVA--SVQNPNAS 320

RESULT 14

ID _RBP1_DROME STANDARD; PRT; 1887 AA.

AC P04052; O9VYX6;

DT 01-NOV-1986 (Rel. 03, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6).

GN RPII215 OR CG1554.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OC NCBI_Taxid=7227;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=89218930; PubMed=2496296;

RA Jokest R.S., Weeks J.R., Zehring W.A., Greenleaf A.L.;

RT "Analysis of the gene encoding the largest subunit of RNA polymerase II in *Drosophila*."

RL Mol. Gen. Genet. 215:266-275(1989).

RN (2)

RP SEQUENCE FROM N.A.

RX STRAIN=Beirley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adamas P.G., Scherer S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amandas P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA Georgia R.A., Lawlis S.R., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abail J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballou R.M., Bau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Bonos P.V., Borman B.P., Bhandari D., Bolashkov S.,

RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Broctler P.,

RA Burla K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cavalay S., Dahlke S., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Maya A.D., Daw I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Pertiera S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,

RA Glodok A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hoatlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mosheiff A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuekern D.R., Paclet J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinart K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,

RA Shue B.C., Siden-Klamow I., Simpson M., Skupski M.P., Smith T.,

RA Slater E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Stizkeas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasmann D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gbha R.A., Myers B.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*."

RL Science 287:2185-2195(2000).

RN (3)

RP SEQUENCE OP 1-27 FROM N.A.

RX MEDLINE=87089662; PubMed=3025586;

RA Searles L.L., Greenleaf A.L., Kemp W.E., Voelker R.A.;

RT "Site of P element insertion and structures of P element deletions in the 5' region of *Drosophila melanogaster* RpiI215."

RL Mol. Cell. Biol. 6:3312-3319(1986).

RN (4)

RP SEQUENCE OF 1-472 FROM N.A.

RX MEDLINE=85282618; PubMed=2992806;

RA Biggs U., Searles L.L., Greenleaf A.L.;

RT "Structure of the eukaryotic transcription apparatus: feature of the gene for the largest subunit of *Drosophila* RNA polymerase II."

RL Cell 42:611-621(1985).

RN (5)

RP SEQUENCE OF 1441-1887 FROM N.A.

RX MEDLINE=88094402; PubMed=3122024;

RA Allison L.A., Wong J.K.-C., Fitzpatrick V.D., Moyle M., Ingles C.J.;

RT "The C-terminal domain of the largest subunit of RNA polymerase II of *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and mammals: a conserved structure with an essential function."

RL Mol. Cell. Biol. 8:321-329(1988).

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS SUBSTRATES.

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + SUBUNIT.

CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM TEN TO TWELVE DIFFERENT POLYPEPTIDES. THE 215 KDA POLYPEPTIDE IS THE LARGEST COMPONENT OF RNA POLYMERASE II.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- PTM: THE TANDEN 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.

CC -1- THE PHOSPHORYLATION ACTIVATES POL2.

CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE III FOR 5S AND TRNA GENES.

CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

CC EMBL; M27431; AAA28868.1; -

CC EMBL; AE003486; AAF8057.1; -

CC EMBL; M14203; AAA28864.1; -

CC EMBL; M1798; AAA28863.1; -

CC EMBL; M19537; AAA28827.1; -

CC PIR; S04457; RNFF2L.

CC FLYBase; FBgn0003277; RPII215.

CC InterPro: IPR000723; RNA_pol_A.

CC InterPro: IPR007081; RNA_pol_Rpb1_1.

CC InterPro: IPR007082; RNA_pol_Rpb1_2.

CC InterPro: IPR007083; RNA_pol_Rpb1_3.

CC InterPro: IPR007084; RNA_pol_Rpb1_4.

CC InterPro: IPR007085; RNA_pol_Rpb1_5.

CC InterPro: IPR007075; RNA_pol_Rpb1_6.

CC InterPro: IPR007073; RNA_pol_Rpb1_7.

CC InterPro: IPR006592; RNA_pol_N.

CC InterPro: IPR000684; RNA_polII_repeat.

CC Pfam; PF04997; RNA_pol_Rpb1_1; 1.

CC Pfam; PF00623; RNA_pol_Rpb1_2; 1.

CC Pfam; PF04983; RNA_pol_Rpb1_3; 1.

CC Pfam; PF05000; RNA_pol_Rpb1_4; 1.

CC Pfam; PF04998; RNA_pol_Rpb1_5; 1.

CC Pfam; PF04992; RNA_pol_Rpb1_6; 1.

CC Pfam; PF04990; RNA_pol_Rpb1_7; 1.

CC Pfam; PF05001; RNA_pol_Rpb1_R; 22.

CC SMART; SMO0663; RPII215; 1.

CC PROSITE; PS00115; RNA_pol_II_REPEAT; 11.

CC Transferrase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat; DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.

CC ZN_FING 67

```

FT DOMAIN 1579 1881 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
FT CONFLICT 319 324 RAMOKS -> GYAKV (IN REF. 4).
FT CONFLICT 450 450 F -> G (IN REF. 4).
FT CONFLICT 455 458 TLAK -> RCTT (IN REF. 4).
FT CONFLICT 461 472 GHRVQLPWS -> VTGESVAST (IN REF. 4).
FT CONFLICT 741 741 R -> H (IN REF. 1).
FT CONFLICT 1485 1524 SMIGAMFVGSGTSPMTPMANCTPRFSPGHV
FT CONFLICT 1506 1508 MTP -> EDSA (IN REF. 1).
FT CONFLICT 1887 1887 D -> DVKGGG (IN REF. 1).
SQ SEQUENCE 1887 AA; 209167 MW; 4EC68C7708A167A3 CRC64;

Query Match 7.1%; Score 124.5; DB 1; Length 1887;
Best Local Similarity 27.1%; Pred. No. 1.4;
Matches 68; Conservative 16; Mismatches 78; Indels 89; Gaps 14;

QY 95 PGLAGTSRVTPTYSQNSPLSARQSP-----IPSYQ-----VSPSSSF----- 136
DB 1591 PG---GASPNVSP---SSPNYSPSPVYASPRVASTNTFNPQSTGVSPPSSGVSPTSPTS 1646
QY 137 -----PSPGRGEPNNMSTPTPPPLKNGCIPSSLSLRI 170
DB 1647 PTVQFQSSPPSPAGSGSNYSPPGNAYSPSSSNYSPPNSPSPTSPTS-----PSS-PSYSP 1701
QY 171 SNSC-PVTTPPVSSPTSKNPKPL-----PMMESIAKQSMIAKQSMASPNYPPFYAVSA 221
DB 1702 TSPCTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 1755
QY 222 PASPTHRHQHTLATIPECDESDS---TVDSGHMISFQKFAQQCFPSASMYPTPTFNL 278
DB 1756 PTPSPYS-----PPSPSYDSPPSPQYTPGS-----PQYSASAKYPTSP---L 1797
QY 279 VKPAQOMSPN 289
DB 1798 YSPSSPQHSPS 1808

RESULT 15
AMNH YEAST STANDARD; PRT; 1367 AA.
AC P08640; P08068;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosylase SI/S2 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-
GN STAI OR STAZ OR MAL5 OR YIR019C.
OC Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota, Fungi, Ascomycota, Saccharomycotina, Saccharomycetes;
OC Saccharomycetales, Saccharomycetaceae, Saccharomycetes.
OX NCBI_TaxID=4932;
RX STRAIN=8288C / AB972;
RP SEQUENCE FROM N.A.
RC Pubmed=8196970;
RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.B., Hornebell T., Hunt S., Jagsels K., Jones M., Lye G.,
RA Moulis S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowlley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RL Nature 387:84-87(1997).
RN 121
RP SEQUENCE OF 1-242 AND 762-1331 FROM N.A.
RX MEDLINE=87194600; Pubmed=3106330;
RA Yamashita I., Nakamura M., Fukui S.;
RT "Gene fusion is a possible mechanism underlying the evolution of
RT STAI.";
RL J. Bacteriol. 169:2142-2149 (1987).
RN 131
RP SEQUENCE OF 1-31 FROM N.A.
RX STRAIN=SPY101-1C;
RK MEDLINE=89031230; Pubmed=3141213;

```

```

RA Bardo J.M., Ianez E., Zalacain M., Claros M.G., Jimenez A.;
RT Similar short elements in the 5' regions of the STA2 and SGA genes
RT from Saccharomyces cerevisiae.";
RL FEBS Lett. 239:119-124 (1988).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
CC glucose residues successively from non-reducing ends of the chains
CC with release of beta-D-glucose.
CC -1- SIMILARITY: TO S.POMBE SPC215.13.
CC -1- SIMILARITY: SOME. TO S.POMBE SPC285.13C.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; 238061; CA86176.1; -
DR EMBL; M16164; AA35014.1; -
DR EMBL; M16165; AA35015.1; -
DR EMBL; X13857; CA32069.1; -
DR PIR; S48478; S48478.
DR SGD; S0001458; MUC1.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0030447; P:filamentous growth; IDA.
DR GO; GO:0007125; P:invasive growth; IMP.
DR GO; GO:0007124; P:pseudohyphal growth; IMP.
KW Hydrolyase; Glycosidase; Polysaccharide degradation; Glycoprotein;
KW Signal; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 1367
FT DOMAIN 210 1367
FT CARBOHYD 817 817 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 874 874 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1367 AA; 136110 MW; 91C0DE2DBD61AA9D CRC64;

Query Match 7.0%; Score 124; DB 1; Length 1367;
Best Local Similarity 27.1%; Pred. No. 1;
Matches 64; Conservative 35; Mismatches 105; Indels 32; Gaps 12;

QY 85 TTYRKGCKPLPGEIAGT---SSRVTPYSQNSPLSARQSP---IPSYQVSPSSSPSPSR 141
DB 698 STTESSAPVPTPSSSTTESSAPVPTPSSSTTESSA---VPP---TPSSSTTESSSA 750
QY 142 GEPNNMSTPTPPPLKNGCIPSSLSLISNCPYTPPVSSPTSKNPKPLN-WEISIAQ 200
DB 751 PVTSTTESSAP-----VPPSSSTTESSAPVPTPSSSTTESSAPVPTPSSSTTES 804
QY 201 SMAIAKQSMASPNYPPFYAVSAPSPTRHQHTLATIPECDESDSTVDSGHMISFQKFA 260
DB 805 SVAPVPTPSSSN---ITSSAPSPPTPSSSTTES-SSVAPVPTPSSSTTESSAPVS-----S 856
QY 261 QQQPPSASVPT-SPTFNLVKPAQOM---SPNTAAFOEIGQS---SEFKENSQVK 310
DB 857 STTESSVAPVPTPSSSNITSSAPSIIPFSSSTTESF-STGTVTPTSSSKYPPSQTE 911

```

Search completed: December 19, 2003, 14:27:05
Job time : 19 secs

QY 66 NEVLKALCVEAGWN

Db 132 NMAAKGL-----TSASLSTKSPFFNSMNVNPLSSQSMSP 167
Qy 123 PIPSYOVSPSSSPSPSPGEPN-----NNMSSTFFPLRNGGIPSSLPLRLISNSCPV 176
Db 168 PMSISMSMSSMSPSAVTGVGSSSLNLSLNNLSLSPSL-NSAVPTP-----ACPY 218
Qy 177 TPPVSSPTSKPKPLPNMESIAKOSMAI-----AKOSMASFNYPFYAVSAPAS 224
Db 219 APPT-----PPYVYRDTCSNLSLRLAKAKQH-SSFGYA--SVQKPAS 258

RESULT 2

US-08-754-477A-121
Sequence 121, Application US/08754477A
Patent No. 6518411
GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC
TITLE OF INVENTION: AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477A
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-754-477A-121

Query Match 7.2%, Score 127.5; DB 4; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.00067;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

Qy 6 ATSTSAIAAAAAAAAAARRKPSWRENNRRRRRAVAIAKIYTLRAQGDYNLRKHCN 65
Db 73 AVMTLTLTEARVAVFKNRRARKRERNOAELCKNGGPGQ-FNGLMOPYDMDYGYSYN 131
Qy 66 NEVLKALCIEAGVVEEDGTTYRKCKPLPGEIATSSRVP-YSSQNSPLS--SAFOS 122
Db 132 NMAAKGL-----TSASLSTKSPFFNSMNVNPLSSQSMSP 167
Qy 123 PIPSYOVSPSSSPSPSPGEPN-----NNMSSTFFPLRNGGIPSSLPLRLISNSCPV 176
Db 168 PMSISMSMSSMSPSAVTGVGSSSLNLSLNNLSLSPSL-NSAVPTP-----ACPY 218
Qy 177 TPPVSSPTSKPKPLPNMESIAKOSMAI-----AKOSMASFNYPFYAVSAPAS 224
Db 219 APPT-----PPYVYRDTCSNLSLRLAKAKQH-SSFGYA--SVQKPAS 258

RESULT 3

US-08-754-477A-125
Sequence 125, Application US/08754477A
Patent No. 6518411
GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
TITLE OF INVENTION: RIEG COMPOSITIONS AND THERAPEUTIC
TITLE OF INVENTION: AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477A
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-754-477A-125

Query Match 7.2%, Score 127.5; DB 4; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.00067;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

Qy 6 ATSTSAIAAAAAAAAAARRKPSWRENNRRRRRAVAIAKIYTLRAQGDYNLRKHCN 65
Db 73 AVMTLTLTEARVAVFKNRRARKRERNOAELCKNGGPGQ-FNGLMOPYDMDYGYSYN 131
Qy 66 NEVLKALCIEAGVVEEDGTTYRKCKPLPGEIATSSRVP-YSSQNSPLS--SAFOS 122
Db 132 NMAAKGL-----TSASLSTKSPFFNSMNVNPLSSQSMSP 167
Qy 123 PIPSYOVSPSSSPSPSPGEPN-----NNMSSTFFPLRNGGIPSSLPLRLISNSCPV 176
Db 168 PMSISMSMSSMSPSAVTGVGSSSLNLSLNNLSLSPSL-NSAVPTP-----ACPY 218
Qy 177 TPPVSSPTSKPKPLPNMESIAKOSMAI-----AKOSMASFNYPFYAVSAPAS 224
Db 219 APPT-----PPYVYRDTCSNLSLRLAKAKQH-SSFGYA--SVQKPAS 258

RESULT 4

US-08-957-351-26
Sequence 26, Application US/08957351
Patent No. 6306586
GENERAL INFORMATION:
APPLICANT: Semina, Elena
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF CATARACTS
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/957,351
FILING DATE: 24-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-024.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-957-351-26

Query Match 7.2%; Score 126.5; DB 4; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.00083;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

QY 6 ATSTSAAAAAAAAAARRKPSWRENNRRRRRAVAIAKIYTGIPAGQDYNLPKHCND 65
DB 73 AWMITLTPARVWFKNRRAKWKRRNQAEICKKGFPGQ-FNGLMQPYDDMYPGYSYN 131
QY 66 NEVLKALCVEAGWVVEEDGTTTRKCKPLPGEIAGTSSRYTP-YSSONOSPLS--SAFOS 122
DB 132 NMAAGL-----TSASLSTSPFPFNSNNVNPPLSSQSNFSP 167
QY 123 PIPSYOVSPSSSPSPSRGEPN-----NNMSSTFPFLRNGGIPSLPSLRISNSCPV 176
DB 168 PMSISMSSSMWVSATVGVGSSLSLNINLNLSSPSL-NSAVPTP-----ACPY 218
QY 177 TPPVSSPTSKNPKPLPNMESIAKOSMAI-----AKOSMASFNPFYVASAPAS 224
DB 219 APPT-----PPYVYRDTGNCSSLASLRLAKQH-SSFGYA--SVONPAS 258

RESULT 5
US-08-754-477A-5
Sequence 5, Application US/08754477A
Patent No. 6518411
GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
APPLICANT: Semina, Elena
TITLE OF INVENTION: RING COMPOSITIONS AND THERAPEUTIC
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/754,477A
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-754-477A-5

Query Match 7.2%; Score 126.5; DB 4; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.00083;
Matches 60; Conservative 25; Mismatches 88; Indels 59; Gaps 11;

QY 6 ATSTSAAAAAAAAAARRKPSWRENNRRRRRAVAIAKIYTGIPAGQDYNLPKHCND 65
DB 73 AWMITLTPARVWFKNRRAKWKRRNQAEICKKGFPGQ-FNGLMQPYDDMYPGYSYN 131
QY 66 NEVLKALCVEAGWVVEEDGTTTRKCKPLPGEIAGTSSRYTP-YSSONOSPLS--SAFOS 122
DB 132 NMAAGL-----TSASLSTSPFPFNSNNVNPPLSSQSNFSP 167
QY 123 PIPSYOVSPSSSPSPSRGEPN-----NNMSSTFPFLRNGGIPSLPSLRISNSCPV 176
DB 168 PMSISMSSSMWVSATVGVGSSLSLNINLNLSSPSL-NSAVPTP-----ACPY 218
QY 177 TPPVSSPTSKNPKPLPNMESIAKOSMAI-----AKOSMASFNPFYVASAPAS 224
DB 219 APPT-----PPYVYRDTGNCSSLASLRLAKQH-SSFGYA--SVONPAS 258

RESULT 6
US-08-754-477A-131
Sequence 131, Application US/08754477A
Patent No. 6518411
GENERAL INFORMATION:
APPLICANT: Murray, Jeffrey
APPLICANT: Semina, Elena
TITLE OF INVENTION: RING COMPOSITIONS AND THERAPEUTIC
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/754,477A
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-022.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 131:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-754-477A-131

Query Match 7.1%; Score 124.5; DB 4; Length 271;
Best Local Similarity 25.9%; Pred. No. 0.0013;
Matches 60; Conservative 24; Mismatches 89; Indels 59; Gaps 11;

QY 6 ATSTAAAAAARRRKPSWRENNRRRRRAVAKIYTLRAQGDYNLPKHCND 65
DB 73 AVMTLTLBARVWPKRPAKWRKBERNOQALCKNGFGPQ-FNGLMDPYDMYGSYN 131
QY 66 NEVLKALCVBAGVVEBDGTTTRKCKPLPGEIAGTSSKATP-YSSQNSPLS-SAFOS 122
DB 132 NMAAKGL-----TSASLSTKSPFFFNMSMNVNPLSSQSMFSP 167
QY 123 PIPSVQSPSSSPSPSKGEPN-----NNMSSTFFPFLRNGGIPSSLPLRISNSCPV 176
DB 168 PMSISMSMSMSWVSATVGVGSSLNLNMLNLSBPL-NSAVPTP-----ACPY 218
QY 177 TPVSSPTSKPKPLPNMESIAKQSMAI---AKQSMASFNYPYAVSAPAS 224
DB 219 APPT-----PPYVVRDTCNSSLSASLRKAKQH-SFVGVA--SVQKPA 258

RESULT 7
US-08-728-323A-2
Sequence 2, Application US/08728323A
Patent No. 5948676

GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: Immediate Early Protein From Kaposi's
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus DNA
TITLE OF INVENTION: Encoding Same And Uses Thereof
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSER: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.10.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,323A
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/52266/JPW/MSC/SKS
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-728-323A-2

Query Match 7.0%; Score 124; DB 2; Length 1162;
Best Local Similarity 22.5%; Pred. No. 0.012;

Matches 79; Conservative 49; Mismatches 129; Indels 94; Gaps 17;

QY 8 STSAAAAAARRRKPSWRENNRRRRRAVAKIYTLRAQGDYNLPKHCND 67
DB 13 STGAPLTRGSCRRKNSPERCDLGLQPRKRVADSI--DGRCEGHTLP----- 63
QY 68 VLKALCVBAGVVEBDGTTTRKCKPLPGEIAGTSSKATP-YSSQNSPLS-SAFOS 118
DB 64 -----IGSPVFTSG---LPAPVSPFLPAPLPSPAPLPALPPLPPT 108
QY 119 AFOSPI-PSYQSPSSSPSPSKRG-EPNNMSSTFFPFLRNGGIPSSLPLRISNSCPV 176
DB 109 TSSSPIPSHPVSGTDTHTSPSPALPPTQSPESQRP-----PLSSPTGRPDSTPM 161
QY 177 TPVSS-----SPTSKPKP--LPNMESIAKQSM-AIKQSMASFNYPF---YAVSAP 222
DB 162 RPPSQQTTPPSPTTPPEPPSKSPSLAPSTRSRKRRLSSPQGSTLNPICOSP 221
QY 223 ASPTRHQHTLATIPEDSDSDSTVDGHWISFQFAQOQPF---SAS-----MVPT 272
DB 222 VSP-----PRCDPANRSV-----YPMATESPIYVGSSSDGDTPPRQPT 261
QY 273 SP-TFNLYKPAPOQMSPTTAFA-----OEIGQSEKFEKNSQVQKPMGE 315
DB 262 SPISGSSSPSGSWGDDTAMLVLAETAEASKVKCKESENQAGEDWD 312

RESULT 8
US-09-298-568-2
Sequence 2, Application US/09298568
Patent No. 6322792

GENERAL INFORMATION:
APPLICANT: Kieff, Elliott D.
APPLICANT: Bailestas, Mary E.
APPLICANT: Kaye, Kenneth M.
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
TITLE OF INVENTION: RHADINO VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
FILE REFERENCE: 16412-10001R
CURRENT APPLICATION NUMBER: US/09/298,568
CURRENT FILING DATE: 1999-04-21
EARLIER APPLICATION NUMBER: US 60/109,422
EARLIER FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 1162
TYPE: PRT

ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-298-568-2

Query Match 7.0%; Score 124; DB 4; Length 1162;
Best Local Similarity 22.5%; Pred. No. 0.012;
Matches 79; Conservative 49; Mismatches 129; Indels 94; Gaps 17;

QY 8 STSAAAAAARRRKPSWRENNRRRRRAVAKIYTLRAQGDYNLPKHCND 67
DB 13 STGAPLTRGSCRRKNSPERCDLGLQPRKRVADSI--DGRCEGHTLP----- 63
QY 68 VLKALCVBAGVVEBDGTTTRKCKPLPGEIAGTSSKATP-YSSQNSPLS-SAFOS 118
DB 64 -----IPSPVFTSG---LPAPVSPFLPAPLPSPAPLPALPPLPPT 108
QY 119 AFOSPI-PSYQSPSSSPSPSKRG-EPNNMSSTFFPFLRNGGIPSSLPLRISNSCPV 176
DB 109 TSSSPIPSHPVSGTDTHTSPSPALPPTQSPESQRP-----PLSSPTGRPDSTPM 161
QY 177 TPVSS-----SPTSKPKP--LPNMESIAKQSM-AIKQSMASFNYPF---YAVSAP 222
DB 162 RPPSQQTTPPSPTTPPEPPSKSPSLAPSTRSRKRRLSSPQGSTLNPICOSP 221
QY 223 ASPTRHQHTLATIPEDSDSDSTVDGHWISFQFAQOQPF---SAS-----MVPT 272
DB 222 VSP-----PRCDPANRSV-----YPMATESPIYVGSSSDGDTPPRQPT 261

Qy 273 SP-TFNLVKPAQOMSPNTAA-----OEIGOSSEFKFENSQVKEWEGE 315
 Db 262 SPISIGSSSSSESGWGDPTMLVLAIETAEASKNKECESENNQAGEDNGD 312

RESULT 9 US-09-410-399-2

Sequence 2, Application US/09410399
 Patent No. 6482587
 GENERAL INFORMATION:
 APPLICANT: Robertson, Eric S.
 APPLICANT: Colter, Murray A.
 TITLE OF INVENTION: Method to Inhibit or Enhance the Binding of Viral DNA
 FILE REFERENCE: US-03778
 CURRENT APPLICATION NUMBER: US/09/410,399
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 2
 LENGTH: 1162
 TYPE: PRT
 ORGANISM: Kaposi's sarcoma-associated herpesvirus
 US-09-410-399-2

Query Match 7.0%; Score 124; DB 4; Length 1162;
 Best Local Similarity 22.5%; Pred. No. 0.012;
 Matches 79; Conservative 49; Mismatches 129; Indels 94; Gaps 17;

Qy 8 STSAAAAAARRRRRRRRRRRRRRAAARAAKYYTLRAGQDYNLPKHDNNE 67
 Db 13 STGAPLTRGSCRRKNSPERCDLGDHLHQPKKHVADSI--DGRCGPTLP----- 63
 Qy 68 VKALCVAEAGWVEEDGTYRKCKPLPGRIAGTSSRTVPYSSONS-----PLSS 118
 Db 64 -----IPGSPVFTSG---LPAFVSSPTLPVAPIPAPATPLPPALLPPT 108
 Qy 119 AFQSPF-PSTYQVSPSSSSFPSPERG-EPNNNSSTFPPLRNGISSLPSLRISNCPV 176
 Db 109 TSSSPFSPHVPVPGTTDHSFSPALPPIQSPSSQRP-----PLSPTGRPDSTM 161
 Qy 177 TPVVS-----SPFSKPKP--LPMESIAKQSM-AIKQSMASNTYF---YAVSAP 222
 Db 162 RPPSPQQTTPHSPPTPPSPSSKSPDLABETLSLRKRRSSQSGSTLNPICOSP 221
 Qy 223 ASPTRHQHTLATIPECDESDSTVDSGHWISFOKFAQOOPF---SAS-----MVP 272
 Db 222 VSP-----PRCDPANRSV-----YPMATESPTLYGSSSDGDTPPROPPT 261
 Qy 273 SP-TFNLVKPAQOMSPNTAA-----OEIGOSSEFKFENSQVKEWEGE 315
 Db 262 SPISIGSSSSSESGWGDPTMLVLAIETAEASKNKECESENNQAGEDNGD 312

RESULT 10 US-08-227-536-2

Sequence 2, Application US/08227536
 Patent No. 5658784
 GENERAL INFORMATION:
 APPLICANT: Eckner, Richard
 APPLICANT: Ewen, Mark
 APPLICANT: Livingston, David
 TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
 TITLE OF INVENTION: FACTOR P300 AND USES OF P300
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
 STREET: Ten Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/227,536
 FILING DATE: 14-APR-1994
 CLASSIFICATION: 436
 ATTORNEY/AGENT INFORMATION:
 NAME: Williams Ph.D., Kathleen A.
 REGISTRATION NUMBER: 34,380
 REFERENCE/DOCKET NUMBER: DFCI-308XX
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 451-0313
 TELEFAX: (617) 542-2290
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2414 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-227-536-2

Query Match 6.9%; Score 122.5; DB 1; Length 2414;
 Best Local Similarity 21.0%; Pred. No. 0.047;
 Matches 94; Conservative 35; Mismatches 166; Indels 153; Gaps 15;

Qy 2 TSDGASTSAAAAAARRRRRRRRRRRRRRAAARAAKYYTLRAGQDYNLPK 61
 Db 594 TDPALAKORRENTLVAVARKVEGDMYESANNRAE-----YHLLAEKIYIK 642
 Qy 62 HCDNN-----EVLKALCVAEAGWV-----VEEDGTYRKCKPLPGEIA 101
 Db 643 ELBEKRTQLQONLPMNAAGVPMNMGQFQPGHT-SNGPLPDSMTIRGSVPNQ 701
 Qy 102 -SSRVTYSSONSQSPSSAFQSPPIPSYQV-----SPSSS 134
 Db 702 NMPRTTPQGLNQFGQMSVAQPIVPRQPPPLQHGGQLAQGALNPMWGPMMQPSNQ 761
 Qy 135 SPSSSRGPPNNMSSTFPPLRNGISSLPSLRISNCPVTPVSPSTK----- 186
 Db 762 GGFPLQTOPFSGQMVNTNIPLASSGQAPVSOAQWSSSSCPVNSPIMPSSQSHIHCPQ 821
 Qy 187 -----NPKPLPMW-----ESIAKQ-----SMIAKQSMASFN 214
 Db 822 LPQALHONSFPVPSKRTPTPHHTPPSIGAQOPATTPAPVPTPPAMPGPQSOALHP 881
 Qy 215 PFYAVSAPSPTRHQHTLATIPECDESDSTVDSGHWISFOKFAQOQPF---ASNV 270
 Db 882 PQTPTPTTQLPQOVQSLAPASADQ-----PQQPRLSQOSTRAVSV 924
 Qy 271 PFSPTFNLVKPA-----QQNSP-----NTAAFOEIGOSSEFKFE--- 305
 Db 925 PT-PNAPLPPQATPLSQPAVISIGQVSNPSTSTEVNSQAIKQSQVEKMAKME 983
 Qy 306 -----NSQVPMGERIHDVGMEDLE 326
 Db 984 VDQPEADTQPEDISESKVEDCKMESTE 1011

RESULT 11 PCT-US95-04682-2

Sequence 2, Application PCT/US9504682
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
 TITLE OF INVENTION: FACTOR P300 AND USES OF P300
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
 STREET: Ten Post Office Square
 CITY: Boston

STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,536
FILING DATE: 14-April-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holliday C. Heine, Ph.D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-308Xq999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 451-0313
FAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04682-2

Query Match 6.9%; Score 122.5; DB 5; Length 2414;
Best Local Similarity 21.0%; Pred. No. 0.047;
Matches 94; Conservative 35; Mismatches 166; Indels 153; Gaps 15;

2 TSDGTSAAATAAAARRRKPSRENNRRRRRAVAAKIYGLRAQGDYNIJPK 61
594 TDPDAALKDRMENLVAAKRVGDMYESANNRAE-----YHLLAEKIKYI 642
62 HCDNN-----EVLKALCEAGWV-----VEEDGTYRKCKPLPGEIAGT 101
643 ELEBRRTTRLOKQNLPRNAAGVPIVSMNGPMGQRPQMT-SNGPLPDSMIRGSVP 701
102 -SSRVTVSSONQSPSSAFQSPISYQV-----SPSSS 134
702 MMRPRTTPOGSLNQFGQMSMAOPPIVPRQTPRLQHGLQALPQALNPMGCGPRMOQPSNQ 761
135 SPSSPSRGEPPNNMSTFFPFLRNGIPLSLRLSNCPYTPPVSSFTSK----- 186
762 GQFLPOTOPPSGGMVNTIPLAPSSGQAFVSOQWSSSSCPVNSPIMPGSGQSHIHCPQ 821
187 -----NPKPLPNW-----ESIAKQ-----SMALAKQSMASFNY 214
822 LPQPLHONSPPVPSRTTTPHTTPPSISAOQPPATTIPAVPTTPPAMPFGQSOALHP 881
215 PFYAVASAPRTHHQFTLATIPEDSDSSSTVDSGHWISFOKFAQQQPFASNV 270
882 PRQTPPTTQAPQVQVPLPAASADQ-----PQQPQSGQSTAAAV 924
271 PTPSPFNLVKAP-----QQNSP-----NTAAFOIGSSSEKPF----- 305
925 PT-PNAPLPPQAPPLSOPAVSIEQVSNPSTSTEVNSQALAEKPSQOEKMAKME 983
306 -----NSQVPMWESERHIDVGMEDLE 326
984 VQDPEDATQPEDESKEVDCMESTE 1011

APPLICANT: CADUS PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: HUMAN MEKX PROTEIN AND NUCLEIC ACID MOLECULES
FILE REFERENCE: CPI-085CPC
CURRENT APPLICATION NUMBER: US/09/423,890
CURRENT FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: USN 60/078,153
PRIOR FILING DATE: 1998-03-16
PRIOR APPLICATION NUMBER: USN 60/099,165
PRIOR FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 8
LENGTH: 1493
TYPE: PRT
ORGANISM: Mus musculus
US-09-423-890-8

Query Match 6.8%; Score 120.5; DB 4; Length 1493;
Best Local Similarity 22.2%; Pred. No. 0.035;
Matches 76; Conservative 36; Mismatches 131; Indels 99; Gaps 12;

5 GATSTAAAAATAAAARRRKPSRENNRRRRRAVAAKIYGLRAQGDYNIJPK 64
139 GATSPAGAEPPSAA-----PSGREMEK-----ETLGLHGMED--RDEEM 179
65 NNEVIALCEAGVVEEDGTYRK-----CKPLPEIAGTSRVTVSSQNSPLSSAFQ 121
180 IREKIKATCMPA-WHEWLERRRNRGRPVVVPPIP--IKDGGSEVNNLAEPQEGAGSA 236
122 SPIPSYQSPSSPSRGRPPNNMSTFFPFLRNGIPLSLRLSNCPY----- 176
237 APAPGRSPSPGSSPSRSVAPES-----BVRKRKRVSPVFGSG 277
177 -TPVSSPTSKNPKPL-PNWEISAKQSMALAKQSMASFNYPFYAVASAPSTHR----- 228
278 RITPPRRAPSPDGFSPVPEETSRVNVKYMRAVLVLQDIGNSFLIGDSFDNKKYRVI 337
229 -----HQPHTLATIPECDS-----STVDS 250
338 GPONSCGRGAFCHILFVLMRLVFOLEPSDPLMKRLKNPEVESLFOKYSRRSSRIYA 397
251 GHWISFOK-----AQQPFASMWPTSPFNLVKPAPQOMSP 288
398 PSRNTIQKVSMSHTLSSSTSTSSSENSIKOEBEOMCP 439

RESULT 13
US-08-628-829-4
Sequence 4, Application US/08628829A
Patent No. 6333170
GENERAL INFORMATION:
APPLICANT: Johnson, Gary L.
TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To Extern
FILE REFERENCE: CPI-004DVC3
CURRENT APPLICATION NUMBER: US/08/628,829A
CURRENT FILING DATE: 1996-04-05
EARLIER APPLICATION NUMBER: 08/440,421
EARLIER FILING DATE: 1995-05-15
EARLIER APPLICATION NUMBER: 08/323,460
EARLIER FILING DATE: 1994-10-14
EARLIER APPLICATION NUMBER: 08/049,254
EARLIER FILING DATE: 1993-05-15
EARLIER APPLICATION NUMBER: 08/410,602
EARLIER FILING DATE: 1995-04-24
EARLIER APPLICATION NUMBER: 08/472,934
EARLIER FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 4
LENGTH: 1593
TYPE: PRT
ORGANISM: Mus musculus

US-08-628-829-4

Query Match 6.8%; Score 120.5; DB 4; Length 1593;

Best Local Similarity 22.2%; Pred. No. 0.039; Mismatches 131; Indels 99; Gaps 12;

Matches 76; Conservative 36; Mismatches 131; Indels 99; Gaps 12;

```
OY 5 GATSTAAAAAARRRKPSWRENNRRRRRAVAKIYTGILRAQGNLPHQCD 64
DB 239 GARSPAGAPPSAA-----PSGRMENK-----ETLKLHKMED--REERN 219
OY 65 NNEVULALVEAGVTVVEEDCTYRKQ--CKPLPSGIAGTSRYTVSSQNGSPPLSARQ 121
DB 280 IREKTKATCMPA-WKHEWLERRRRRRGPVVKPIP--IKGDSEVNNMLAAPQEGQAGSA 336
OY 122 SPIPSYQVSPSSSPSPSGEENNNMSTFPPLNNGIPLSLPLRLINSQPV----- 176
DB 337 AAPAPKRRSPSPSPSGRSVKPES-----PQVRRKRVSPVPQSG 377
OY 177 --TPVVSPTSKNPKPL-PNWSIAKQMAIAKQNASPVYFVAVSAPSPTR----- 228
DB 376 RITPRRAPSDFGFSYSPSESTRNVKYMARLYLLOQIGPNSFLIGDSDPNKTRVFI 437
OY 229 -----HOFHTLATIPECDESD-----SSTVDS 250
DB 438 GPQNGCGGACFCIHLLFVMLRVQLPQLEPQDPMLKTLKNPEVESLQKYHSRRSRKA 497
OY 251 GHWISFOK-----AQOQPSASVNPSTFNLVKPAPQOMSP 288
DB 498 PSRNTIOKRVSRMSNSHTLSSSTSTSSSENSIKDEEQWCP 539
```

RESULT 14

US-08-040-548-4

Sequence 4, Application US/08040548

Patent No. 5763209

GENERAL INFORMATION:

APPLICANT: Sukhatme, Vikas P.

TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE

NUMBER OF SEQUENCES: 67

TITLE OF INVENTION: FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS

CORRESPONDENCE ADDRESS:

ADDRESS: Arnold, White & Durkee

STREET: 321 No. 5763209th Clark Street, Suite 800

CITY: Chicago

STATE: Illinois

COUNTRY: U.S.A.

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/040,548

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Coughlin, Daniel F.

REGISTRATION NUMBER: 36,111

REFERENCE/DOCKET NUMBER: arcd067

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 744-0090

TELEFAX: (312) 245-4961

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 114 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-040-548-4

Query Match 6.6%; Score 116; DB 1; Length 114;

Best Local Similarity 31.5%; Pred. No. 0.0022;

Matches 29; Conservative 22; Mismatches 27; Indels 14; Gaps 3;

```
OY 102 SSRVTYSSQNGSPPLSAFQSPPIPSYQVSPSSSPSPSPR-GEPNMMSSTFPPLRNGC 160
DB 19 SSYSPVATYSPSPATSPFPSPVPTYSYSPGSSSTYSPAHSGFPPSVATTF-----AS 72
OY 161 IP-----SSPLRLINSQCVTPPVSSPTS 185
DB 73 VPPAFPTQVSPFSPSAGVSSSFSTGLSDMTA 104
```

RESULT 15

US-08-466-344-4

Sequence 4, Application US/08466344

Patent No. 5773583

GENERAL INFORMATION:

APPLICANT: Sukhatme, Vikas P.

TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE

NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS:

ADDRESS: Arnold, White & Durkee

STREET: 321 No. 5773583th Clark Street, Suite 800

CITY: Chicago

STATE: Illinois

COUNTRY: U.S.A.

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,344

FILING DATE: 06-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/040,548

FILING DATE: 31-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Coughlin, Daniel F.

REGISTRATION NUMBER: 36,111

REFERENCE/DOCKET NUMBER: arcd067

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 744-0090

TELEFAX: (312) 245-4961

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 114 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-466-344-4

Query Match 6.6%; Score 116; DB 1; Length 114;

Best Local Similarity 31.5%; Pred. No. 0.0022;

Matches 29; Conservative 22; Mismatches 27; Indels 14; Gaps 3;

```
OY 102 SSRVTYSSQNGSPPLSAFQSPPIPSYQVSPSSSPSPSPR-GEPNMMSSTFPPLRNGC 160
DB 19 SSYSPVATYSPSPATSPFPSPVPTYSYSPGSSSTYSPAHSGFPPSVATTF-----AS 72
OY 161 IP-----SSPLRLINSQCVTPPVSSPTS 185
DB 73 VPPAFPTQVSPFSPSAGVSSSFSTGLSDMTA 104
```

Search completed: December 19, 2003, 14:27:39

Job time : 23 secs


```

Db      823 TCCATGGCCATTGCTAAACATCAATGGCGCTTTTAAATTCCTTTCTATGGCGTTCT 882
Oy      221 AAlProAlaSerProThrHisArgHisGlnPheHisThrLeuAlaThrIleProGluCys 240
      883 GCACTGCTAGTCCGACACATCCGACAGTTTCATACCCGGCTACTATACCTGAAATG 942
Oy      241 AApGluSerAspSerSerThrValaAspSerGlyHisTyrIleSerPheGlnLysPheAla 260
      943 GATGAGCTGACCTTCCACTGTTGATTTGCTGATTCGATTAAGCTTTTCAGAACTTTGCA 1002
Oy      261 GlnGlnGlnProPheSerAlaSerMetValProThrSerProThrPheAsnLeuValLys 280
      1003 CAACAACACCCATCTCTCTCTCTATGCGCAACCTCTCCATCTTCAATCTTGGA 1062
Oy      281 ProAlaProGlnGlnMetSerProAsnThrAlaAlaPheGlnGluIleGlyGlnSerSer 300
      1063 CTGGGCTTCAACAGATGCTCCAAATACCTGCTGCTTCCAAAGATGGTCCAAACCTCT 1122
Oy      301 GluPheLysPheGlnLysAsnSerGlnValLysProTyrGluGlyValLysArgIleHisAspVal 320
      1123 GAGTTTAATTTGAGAAATGCCAAGTTAAACCTGGGAGAGAGATACATGATGTG 1182
Oy      321 GlyMetGluAspLeuGlnLeuThrLeuGlyAsnGlyLysAlaArgGly 336
      1183 GGTATGAGAGATCTTGAGCTTACACTTGGAATGGGAAGGCTCGTGT 1230

RESULT 3
US-09-995-938a-2
/ Sequence 2, Application US/0995938A
/ Publication No. US20030150026A1
/ GENERAL INFORMATION:
/ APPLICANT: JOANNE CHORY
/ APPLICANT: ZHIYONG WANG
/ TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
/ TITLE OF INVENTION: HORMONE ACTION IN PLANTS
/ FILE REFERENCE: SALKINS.046A
/ CURRENT APPLICATION NUMBER: US/09/995, 938A
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 2687
/ TYPE: DNA
/ ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938a-2

Alignment Scores:
Pred. No.: 1,226-139 Length: 2687
Score: 1671.00 Matches: 335
Percent Similarity: 74.28% Conservative: 0
Best Local Similarity: 74.28% Mismatches: 1
Query Match: 94.67% Indels: 116
DB: 13 Gaps: 1

US-09-995-938a-7 (1-336) x US-09-995-938a-2 (1-2687)
Oy      1 MetThrSerAspGlyAlaThrSerThrSerAlaAlaAlaAlaAlaAlaAlaAla 20
      1103 ATGACTTCGATGAGTACGTCACATCAGCAGCTGCGGCGCGGCGGCGGCGGCGGCGG 1162
Oy      21 AlaArgArgLysProSerTyrArgGluArgGluAsnAsnArgArgArgGluArgArg 40
      1163 GCGAGGAGGAGGAGCGCTCGTGAGAGAAAGGAGAAATTCGAGAGAGAGAGAGAGAGAG 1222
Oy      41 ArgAlaValAlaAlaLysIleTyrThrGlyLeuArgAlaGlnGlyAspTyrAsnLeuPro 60
      1223 AGACCTGAGCTGCGCAATATACCTGCGCTTAAGCTCAAGGTGATTAATTAATTTGCT 1282
Oy      61 LysHisCysAspAsnAsnGluValLeuLysAlaLeuCysValGluAlaGlyTyrValAla 80
      1283 AAACATGTGTATTAATTAATGAAGCTCTTTGTGTGAAGCTGCTGGTTGTTT 1342
Oy      81 GluGluAspGlyThrThrTyrArgLys----- 89

```

```

Db      1343 GAAGAAGATGATCTACTATATGCAA-GGTGAAGACTTTCTCATTTTTCAGATCTGA 1401
Oy      89 ----- 89
Db      1402 GCTTTTATTTGATCTTTTGTATGTTGAATCTGAATTCGTTGATTTCAATTGTGTTA 1461
Oy      89 ----- 89
Db      1462 AATGGGTTGAATCTGAGAAATTTGAGGGTTTCTCAAGATGAATTTGAATCATCAGAAAC 1521
Oy      89 ----- 89
Db      1522 TATGATGATCTGATTTCTCAAGATGAATTTATG-3GTTTCTTAATTTTGAAGTTA 1581
Oy      89 ----- 89
Db      1582 TTATGTGATGCTAAAGCTTAATCTTTATGTATATCTTGCTCAAGATCATTTGCAT 1641
Oy      89 ----- 89
Db      1642 TGTGTTTCTTTGCTTACCTGTGATGATGATGTTGATGTGTTATTTGTTGCTTTTG 1701
Oy      90 ----- 105
      1702 TTGAGATATCAGGAGATGCAAGGCTTTTACCTGTGATGATGCTGGGAGCTTCATCTCAGTA 1761
Oy      106 ThrProTyrSerSerGlyAsnGlnSerProLeuSerSerAlaPheGlnSerProIlePro 125
      1762 ACTCATATTCTATCAGACAGACAGACCTCTTCTTCACGCTTCAAGTCCATCCCA 1821
Oy      126 SerTyrGlnValSerProSerSerSerSerPheProSerProSerArgGlyGluProAsn 145
      1822 TCTTACCAAGTTAGGCGGCTTCTTCATCATTCGCCAGCTCTTCGCGGTGAACCAAT 1881
Oy      146 AsnAsnMetSerSerThrPhePheProPheLeuArgAsnGlyLysIleProSerSerLeu 165
      1882 AACACATGTCCTCTACATTTCTTCTCTTCCAGAAATGGTGCACTTCTTCTCTT 1941
Oy      166 ProSerLeuArgIleSerAsnSerCysProValThrProProValSerSerProThrSer 185
      1942 CTTTCTCCAGATTTCAACAGATTTGTCAGTTACCCAGCTTCATCGCGCATCTTCT 2001
Oy      186 LysAsnProLysProLeuProAsnTyrPheGlnSerIleAlaLysGlnSerMetAlaIleAla 205
      2002 AAGAACCCGGAAGCGTTGCTTACCTGAGATATATCGTACCAATCCATGGCCATTTGCT 2061
Oy      206 LysGlnSerMetAlaSerPheAsnTyrProPheTyrAlaValSerAlaProAlaSerPro 225
      2062 AAACATCAATGGCGCTTTTAATTAATCTTTCTATGGGGTTTCGACCTGCTACTCGG 2121
Oy      226 ThrHisArgHisGlnPheHisThrLeuAlaThrIleProGluCysAspGluSerAspSer 245
      2122 ACACATGCGCACAGTTTCATACCTGCTACTATACCTGAATGTGATGAGTCTGACTCT 2181
Oy      246 SerThrValaAspSerGlyHisTyrIleSerPheGlnLysPheAlaGlnGlnGlnProPhe 265
      2182 TCCACTGTGATTTCTGCGATTCGATTAAGCTTTTCGAAGTTTGCCACACACAGCCATTC 2241
Oy      266 SerAlaSerMetValProThrSerProThrPheAsnLeuValLysProAlaProGlnGln 285
      2242 TCTGCTCTATGAGGCCAACCTCTCTCACTTCAATCTTTGTGAACCTGCGCTCGAGAG 2301
Oy      286 MetSerProAsnThrAlaAlaPheGlnGluIleGlyGlnSerSerGluPheLysPheGlu 305
      2302 ATGTCTCAAAATACCTGCTCTTCCAAAGATGTGCTCAAGCTCTGAGATTTAAATTTGAG 2361
Oy      306 AsnSerGlnValLysProTyrGluGlyValLysArgIleHisAspValAlaMetGluAspLeu 325
      2362 AATAGCCAAATTAACCTGGGAGAGAGAGATACATGATGTGGATGAGGATCTT 2421
Oy      326 GluLeuThrLeuGlyAsnGlyLysAlaArgGly 336

```

Db 2422 GAGCTTACCTTGGAAATGGGAGGCTCGTGT 2454

RESULT 4
US-09-995-938A-1
Sequence 1, Application US/0995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
APPLICANT: ZHIYONG WANG
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
FILE REFERENCE: SALKINS 046A
CURRENT APPLICATION NUMBER: US/09/995, 938A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2687
TYPE: DNA
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-1

Alignment Scores:
Pred. No.: 5,15e-139 Length: 2687
Score: 1664.00 Matches: 334
Percent Similarity: 74.06% Conservative: 0
Best Local Similarity: 74.06% Mismatches: 2
Query Match: 94.28% Indels: 116
Gaps: 1

US-09-995-938A-7 (1-336) x US-09-995-938A-1 (1-2687)

Qy 1 MetThrsEsrApgIyAlaThrSerThrsEsrAlaAlaAlaAlaAlaAlaAla 20
Db 1103 ATGACTTCGATGGAGCTACGTCGACATCAGCAGCTGCGCGCGGACAGCAGC 1162

Qy 21 AAlaArgAlvLpProSerTTPArgLlUArgLlUaMaMaArgArgArgArgArg 40
Db 1163 GCGGAGGAGGAGCCGCTCGGAGAGAGGAGGAGTAATCGGAGAGAGAGAGGAG 1222

Qy 41 ArgAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 60
Db 1223 AGAGCTGAGCTGCGAGAGATATACCTGAGCTTACAGCTCAAGGATTAATTTGCT 1282

Qy 61 LysHisCyAspAspAsnGluValLeuLysAlaLeuCyValGluAlaGlyTyrValVal 80
Db 1283 AAACATTTGATTAATGAAGACTCTTAAGCTCTTTGTTGAAGCTGGTTGGTTGT 1342

Qy 81 GluGluAspGlyThrThrTyrArgLys----- 89
Db 1343 GAAGAAAGATGGTACTACTTATGCAA-GGTGAAGACTTTCATTTTTCAGATCTGA 1401

Qy 89 ----- 89
Db 1402 GCTTGTATTATGATGTTTGTGATGTTGAATCTGAATTCGTTGATTCATTTGGTTA 1461

Qy 89 ----- 89
Db 1462 AATGGTTGAATCTGAGAAATTTGAGGGTTTCTCAAGTGAATTTGATCATCAGAAAC 1521

Qy 89 ----- 89
Db 1522 TATGATGATCTGATTTCTCAAGTGAATTTATGGGTTTCTTTCTAATTTAGAGTTA 1581

Qy 89 ----- 89
Db 1582 TTATTGATAGCTAAAGCTTAATCTTTATGATGATGATCTTGTCGAAGCAATTCAT 1641

Qy 89 ----- 89
Db 1642 TGTGTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1701

Qy 90 -----GlyCylAspProLeuProGlyGluLeuAlaGlyThrSerArgVal 105

Db 1702 TTGGAGTATCAGGATGATCAGCCCTTTAATCTGTAGATAGCTGGAGACTTCTCGAGTA 1761

Qy 106 ThrProTyrSerSerGlnAsnGlnSerProLeuSerSerAlaPheGlnSerProLeu 125
Db 1762 ACTCATATTCATCATCAGAAACAGAGCCCTTTATCATGAGCTTTCAAAAGTCCATCCCA 1821

Qy 126 SerTyrGlnValSerProSerSerSerPheProSerProSerArgGlyGluProbn 145
Db 1822 TCTTACCAAGTATCCCGCTCTTTCATCATCTCCGAGCTCTTCTCGGCGTGAACCAAT 1881

Qy 146 AsnAsnMetSerSerThrPhePheProPheLeuValGlnGlyLysLeuProSerSerLeu 165
Db 1882 AACCAATGCTCTTATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1941

Qy 166 ProSerLeuArgLysSerAsnSerCybProValThrProProValSerSerProThrSer 185
Db 1942 CTTTCCCTCAGAAATCTCAACAGTTGTCAGTTACCCACCGGCTCATCCCGCATCTCT 2001

Qy 186 LysAsnProLysProLeuProAsnTyrGluSerLysAlaLysGlnSerMetAlaLeu 205
Db 2002 AAGAACCCGAACCGTTCCTTACCTGAGATCTATGCTAAGCAATCCATGCGCATTTGCT 2061

Qy 206 LysGlnSerMetAlaSerPheAsnTyrProPheTyrAlaValSerAlaProAlaSerPro 225
Db 2062 AAACCAATCAATGAGCTCTTTTATATATCTTCTTATGCGGTTCTGCACTGCTATGCTCG 2121

Qy 226 ThrHisArgHisGlnPheHisThrLeuAlaThrLysProGlyCyAspGlySerAspSer 245
Db 2122 ACACATGCGCAGAGTTTCAATCACCCTGCTACTATACCTGATGATGAGTCACTACT 2181

Qy 246 SerThrValAspSerGlyHisTyrLysPheGlnLysPheAlaGlnGlnLysProPhe 265
Db 2182 TCCACTGTGATTTGTGTCATTTGATTAAGCTTTTCAAGAGTTTGCAACACAGCATTC 2241

Qy 266 SerLysSerMetValProThrSerProThrPheAsnLeuValLysProAlaProGlnGln 285
Db 2242 TCTGCTCTTATGAGCCCAACTCTTCTTCAATCTTTCATCTTTCGAAACCTGCGCTCAGCG 2301

Qy 286 MetSerProAsnThrAlaAlaPheGlnGlnLysGlnSerSerGluPheLysPheGln 305
Db 2302 ATGCTCCAAATATAGCTGCTCTTCCAAAGATTTGTTCAAGCTCTGAGTTTAAATTTGAG 2361

Qy 306 AsnSerGlnValLysProTyrGluGlyLysArgLysHisAspValGlyMetGluAspLeu 325
Db 2362 AATAGCCCAAGTTAAACCTCGGAGAGAGAGATGATGATGATGATGATGATGAT 2421

Qy 326 GluLeuThrLeuGlyAsnGlyLysAlaArgGly 336
Db 2422 GAGCTTACCTTGGAAATGGGAGGCTCGTGT 2454

RESULT 5
US-09-995-938A-12
Sequence 12, Application US/0995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
APPLICANT: ZHIYONG WANG
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
FILE REFERENCE: SALKINS 046A
CURRENT APPLICATION NUMBER: US/09/995, 938A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 1213
TYPE: DNA
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-12

Alignment Scores:
Pred. No.: 3,67e-127 Length: 1213


```

Qy 200 GlnSerMetAlaIle---AlaLysGlnSerMetAlaSerPheAsnTyrProPheTyrAla 218
Db 738 CAATTCATGCTCATGGCTGCTTAAACAGTCAATGACTCTTTGAACCTCCGTTTATGCG 797
Qy 219 ValSerAlaProAlaSerProThrHisArgHisGlnPheHisThrLeuAlaThrIlePro 238
Db 798 GTGCTGCAGCTCCGACGCTCACTCATGCGCAGTTCCATGCTCCGCTCACTTACTACT 857
Qy 239 GluCyAspGluSerAspSerSerThrValAspSerGlyHisTyrIleSerPheGlnLys 258
Db 858 GAATGTGATGAGTCACTCTTCCACTGTGATTCGTGATTCGATTAAGCTTTCAAAAG 917
Qy 259 PheAlaGlnGlnGlnProPheSerAlaSerMetValProThrSerProThrPheAsnLeu 278
Db 918 TTGACACACACACACCACTCTCTCTCTATGTCACCACTCCGCTCACTTCAATCTC 977
Qy 279 ValLysProAlaProGlnGlnMetSerProAsnThrAlaAlaPheGlnGlnIleGln 298
Db 978 GTGAACCTGCACCAACCAATTCCTCCAAACACAGCAGCAATCCAAAGATTGCTCA 1037
Qy 299 SerSerGluPheLysPheGluAsnSerGlnValLysProTyrGluGlnIleGlnHis 318
Db 1038 AGCTCCGAGTTTAAGTTTGAGAACAGCCAAATTAAAGCCATGGAGAGGAGAGATCCAT 1097
Qy 319 AspValGlyMetGluAspLeuGluLeuThrLeuGlyAsnGlyLysAla 334
Db 1098 GATGGCTATGAGAGATCTAGAGCTCAGCTTGAAATGTAAAGCT 1145

```

RESULT 7

US-09-995-938a-9

Sequence 9, Application US/09995938A

Publication No. US20030150026A1

GENERAL INFORMATION:

APPLICANT: JOANNE CHORY

APPLICANT: ZHIYONG WANG

TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID

TITLE OF INVENTION: HORMONE ACTION IN PLANTS

FILE REFERENCE: SALKINS.046A

CURRENT APPLICATION NUMBER: US/09/995,938A

NUMBER OF SEQ ID NOS: 14

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 9

LENGTH: 3000

TYPE: DNA

ORGANISM: ARABIDOPSIS THALIANA

US-09-995-938a-9

Alignment Scores:

```

Pred. No.: 7,62e-119 Length: 3000
Score: 1439.00 Matches: 298
Percent Similarity: 69.68% Conservative: 10
Best Local Similarity: 81.42% Mismatches: 23
Query Match: 81.53% Indels: 112
DB: 13 Gaps: 6

```

US-09-995-938a-7 (1-336) x US-09-995-938a-9 (1-3000)

```

Qy 1 MetThrSerAspGlyAlaThrSerThrSerAlaAlaAlaAlaAlaAlaAla 20
Db 1337 ATGAGCTCTGACGAGCAAGCTGAGCTCA--GCTGACAGCTGCACAGCAGGATGGCG 1393
Qy 21 AlaAspGlyLysProSerTyrArgGluArgGluAsnAspArgArgGluArgArg 40
Db 1394 ACGAGAGAGAAACCGTCGTGAGAGAGAGAGAGAAACATCGAGAGAGCGCGGAGA 1453
Qy 41 ArgAlaValAlaAlaLysIleTyrThrGlyLeuArgAlaGlnGlyAspTyrAsnLeuPro 60
Db 1454 AGAGCTGTTCGCGGAGATTTATCTGCTTGAAGCTCAAGTAACTAATCTTTCCA 1513
Qy 61 LysHisCysAspAsnGlnLysValLysValLeuCysValGlnIleGlyTyrValVal 80
Db 1514 AAACATTTGACAAACAATGAGTTCTTAAGGCTCTTTGTTCTGAAGCTGGTGGTTGTT 1573

```

```

Qy 81 GluGluAspGlyThrThrTyrArgLysGly----- 90
Db 1574 GAAGAGAGGAGAACTTATATGCA--GATCAGTATCAAAAGCATTTTACTTAGATCTG 1632
Qy 90 ----- 90
Db 1633 ATGTTATCTGAATGATTAGCTGTTGATCTGAAGATTGGATTGTAATTGGTCAATG 1692
Qy 90 ----- 90
Db 1693 GGATTTCTTGAGCTATGATTCAGAGTTTCTAGTACGAGAGAGTCTAGTTTATTCTTAAA 1752
Qy 90 ----- 90
Db 1753 TTGATGAGANTCTTCGCGAAGAAAGTACCTTGAAGGTTCTTCTTAAATTGAGAA 1812
Qy 90 ----- 90
Db 1813 CCGAATTAGCTTACTTCACTTGGTACTATATTAGATCTCTCTTACTTACTTGGATT 1872
Qy 91 ---Cys----- 1932
Db 1873 GATGTGACATGTGATGTTTGTGATTTGTTCTATGAGCAACAGGAGACAAAGCTCTTA 1932
Qy 95 ProGlyGluIleAlaGlyThrSerSerArgValThrProTyrSerSerGlnAsnGlnSer 114
Db 1933 CCGTGAGATGGAGATCACTTCTCGAGCAATCTCTTACTCTTCCATTAACAAAGT 1992
Qy 115 ProLeuSerSerAlaPheGlnSerProIleProSerTyrGlnValSerProSerSerSer 134
Db 1993 CCTCTTTCTTCCACTTTGATAGCCCATCTTATCTTACCAAGTCACTCTCTCTTCT 2052
Qy 135 SerPheProSerProSerArg---GlyLysProAsnAspAsnMetSerSerThrPhePhe 153
Db 2053 TCATCCGAGACCTTCGAGTGTGATGATGCACACAAAT-----ATCTCCAAATCTTC 2106
Qy 154 ProPheLeuArgAsnGlyGlyIleProSerSerLeuProSerLeuArgIleSerAsnSer 173
Db 2107 CCTTTCCTCAGAAAGGAGGATTCCTTCATGCTTCCATCTTCAATGAATCTCAACAGT 2166
Qy 174 CysProValThrProProValSerSerProThrSerLysAsnProLysProLeuProAsn 193
Db 2167 GCTCTGTACCTCCACAGTGTCAATCCCACTTCTAGAAACCCAAACCATTTGCTACT 2226
Qy 194 TrpGluSerIleAlaLysGlnSerMetAlaIle---AlaLysGlnSerMetAlaSerPhe 212
Db 2227 TGGGAATCTTTAACCAACCAATCCATGTCATGCTGTAACAGTCAATGACTTCTTGG 2286
Qy 213 AsnTyrProPheTyrAlaValSerAlaProAlaSerProThrHisArgHisGlnPheHis 232
Db 2287 AACTACCCCTTTTATGCGGTGTCTGCACTGCGCAATCTCACTCATATCGCCAGTTCCAT 2346
Qy 233 ThrLeuAlaThrIleProGluCysAspGluSerAspSerSerThrValAspSerGlyHis 252
Db 2347 GCTTGTGCTACATACCTGAATGTAGTGTGACTTTCACCTGTTATTTGTGTCAT 2406
Qy 253 TrpIleSerPheGlnLysPheAlaGlnGlnGlnProPheSerAlaSerMetValProThr 272
Db 2407 TGGATAGCTTCAAAAGTTTCCACAAACAGCCATTTCTGCTCTATGAGGCCAAC 2466
Qy 273 SerProThrPheAsnLeuValLysProAlaProGlnGlnMetSerProAsnThrAlaAla 292
Db 2467 TCGCTTACCTTCAATCTCGTGAACCTGCACCAACCAATTTCTCCAAACACAGCGCA 2556
Qy 293 PheGlnGluIleGlyGlnSerSerGluPheLysPheGluAsnSerGlnValLysProTyr 312
Db 2527 ATCCAGAGATTTGCTAAGCTCCAGTTTAACTTTGAGAACCCCAAGTTAAGCCATGG 2586
Qy 313 GluGlyGluArgIleHisAspValGlyMetGluAspLeuGluLeuThrLeuGlyAsnGly 332
Db 2587 GAAGGAGAGAGATCATGATGTGCTATGAGAGATCTAGAGCTCAAGCTTGAATAATGAT 2646

```

QY 333 LysAla 334
Db 2647 AAGCT 2652

RESULT 8
US-09-995-938a-3
Sequence 3, Application US/09995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
APPLICANT: ZHIYONG WANG
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
TITLE OF INVENTION: HORMONE ACTION IN PLANTS
FILE REFERENCE: SALKINS.046A
CURRENT APPLICATION NUMBER: US/09/995.938A
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 3000
TYPE: DNA
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938a-3

Alignment Scores:
Pred. No.: 3,22e-118 Length: 3000
Score: 1432.00 Matches: 297
Percent Similarity: 69.46% Conservaive: 10
Best Local Similarity: 67.19% Mismatches: 24
Query Match: 81.13% Indels: 112
DB: 13 Gaps: 6

US-09-995-938a-7 (1-336) x US-09-995-938a-3 (1-3000)

QY 1 MetThSerhPpGlyAlaThrSerThSerAlaAlaAlaAlaAlaAlaAla 20
Db 1337 ATGACGCTGACGAGAGCAAGTCGACCTCA---GCTCAGCTGACAGCAGCATGGCG 1393

QY 21 AlaArgArgLysProSerTrpArgGluArgGluArgGluArgGluArgGluArg 40
Db 1394 ACGAGAGAGGAAACCGCTGCTGAGAGAGAGAGAAACATCGAGAGAGAGCGCGAGA 1453

QY 41 ArgAlaValAlaAlaLysIleTyrThrGlyLeuArgGluGlnGlyAspTyrAsnLeuPro 60
Db 1454 AGCGAGCTGCGCGAGAGATTACTGCTTACGCTCAAGCTCAAGTCAATCTTCCA 1513

QY 61 LysHisCysAspAsnGluValLeuLysAlaLeuCysValGluAlaGlyTrpValVal 80
Db 1514 AACCATTTGTACAAACATGAGGTTCTTAAGGCTCTTTGTTCTGAACTGGTGGTGTT 1573

QY 81 GluGluAspGlyThrTyrArgLysGly----- 90
Db 1574 GAAGAGACGAGAACTACTATCGCAA-GGTCAAGTCAAAAGCATTTTACTAGATCTG 1632

QY 90 ----- 90
Db 1633 AAGTTATCTGATGATTAGCTGTGAATCGAATTGGAATTGCAATTG 1692

QY 90 ----- 90
Db 1693 GGATTTCTTGCTATGAATTCGAGGTTTACTGAGAGAGCTCAAGTTTATTCTAAAA 1752

QY 90 ----- 90
Db 1753 TTGATCGAGATTCTTGGGAGAAAGTACCTTTAGGCTTCTTCTTACTAATTGAGAA 1812

QY 90 ----- 90
Db 1813 CGGAATTAGCTTACTTCACTGTTACTATATTAGATCTCTCTTACTTTGATT 1872

QY 91 ---Cys-----LysProLeu 94
Db 1873 GATTTGACATTTGATGTTTGGTATTTGTTCTATGAGCAAGGAGCAAGAGCTCTTA 1932

QY 95 ProGlyGluIleAlaGlyThrSerSerArgValThrProTyrSerSerGlnAsnGlnSer 114
Db 1933 CCGTGATACATGGCTGGATCATCTTCTCGACCAATCTCTTACTCTTCCATACCAAACT 1992

QY 115 ProLeuSerSerAlaPheGlnSerProIleProSerTyrGlnValSerProSerSerSer 134
Db 1993 CTTCTTCTTCACTTTGATGAGCCCATCTTATTATACCAAGTCAGTCCTTCTCTCTCT 2052

QY 135 SerPheProSerProSerArg---GlyAluProAsnAsnAsnMetSerSerThrPhePhe 153
Db 2053 TCATTCGAGACTCTTCTCGAGTGTGATCACAACAT-----ATTCACAATCTTC 2106

QY 154 ProPheLeuArgAsnGlyIleProSerSerLeuProSerLeuArgIleSerAsnSer 173
Db 2107 CTTTCCCTCAGAGATGGTGATTCCTTCACTGCTTCCATCATTAGATCTCAAAACGT 2166

QY 174 CysProValThrProProValSerSerProThrSerTyrAsnProLysProLeuProAsn 193
Db 2167 GCTCTGTCACTCCACAGTGTCACTCCCACTTTAGAAACCCCAACCATTTGCTTACT 2226

QY 194 TrpGluSerIleAlaLysGlnSerMetAlaIle---AlaLysGlnSerMetAlaSerPhe 212
Db 2227 TGGGAATCTTTACCAACAAATCCATGCTCATGCTCTTAACAGTCAATGCTTCTTG 2286

QY 213 AsnTyrProPheTyrAlaValSerAlaProAlaSerProThrHisArgHisGlnPheHis 232
Db 2287 AACTACCCGTTTATGCGGTGTGCACCTGCCAATCTTACTATCATCGCCAGTTCCAT 2346

QY 233 ThrLeuAlaThrIleProGluCysAspGluSerAspSerSerThrValAspSerGlyHis 252
Db 2347 GCTCCGGCTACTATCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2406

QY 253 TrpIleSerPheGlnLysPheAlaGlnGlnPheSerAlaSerMetValProThr 272
Db 2407 TGAATAGCTTTAAAGTTTGCACAAACAGCAATCTCTGCTTATGCTGCAACCC 2466

QY 273 SerProThrPheAsnLeuValLysProAlaProGlnGlnMetSerProAsnThrAlaAla 292
Db 2467 TCGCATACCTTCAATCTCGTGAACCTGCACACAGCAATGCTTCCAAACAGCAGCA 2526

QY 293 PheGlnGluIleGlyGlnSerSerGluPheLysPheGluAsnSerGlnValLysProTrp 312
Db 2527 ATCCAAAGATTGCTCAAACTCCGAGTTTAACTTGAAGAAACCAAGTTAAAGCATG 2586

QY 313 GluGlyGluArgLysIleAspValGlyMetGluAspLeuGluLeuThrLeuGlyAsnGly 332
Db 2587 GAAGGGAGAGATCCATGATGTGGCTATGAGAGATCTAGAGCTCAGCGTTGAAATGCT 2646

QY 333 LysAla 334
Db 2647 AAGCT 2652

RESULT 9
US-09-938-842A-365
Sequence 365, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCLIP130-3
CURRENT APPLICATION NUMBER: US/09/938.842A
PRIORITY FILING DATE: 2001-08-24
PRIORITY FILING DATE: 2000-08-24
PRIORITY FILING DATE: 2000-08-24
PRIORITY FILING DATE: 2001-01-16
PRIORITY FILING DATE: 2001-06-22

[illegible][illegible]


```

QY 65 -----AenAenGluValIleuLysAlaLeuGlyValGluAlaGlyTrpValValGI 81
DB 1174521 CCACCCCTCCACGACGACACTCTGCACTGCTG----- 1174488
QY 81 uGluAepGlyThrThrTyrrArgLysGlyCyAspProLeuPro-----GlyGluI 98
DB 1174487 -GGAGACATCACCCCCCAGCCGTCACACATCCGCTTCTTCCACCGCTCTAGGGACCT 1174429
QY 98 lAlaGlyThrSerSerArgValThrPro-----TyrSerSerGlnAenGlnSerProL 116
DB 1174428 GGCTGTGTACACCAACACCTGGACCGCCGCTACTGGTACCGCAACTCCACACGCGCT 1174369
QY 116 euSerSerAlaPheGlnSer-----ProIleProSerTyrGlnValSerProSerSerS 135
DB 1174368 CCGCTTCACCCACGACGACCTGACGCGGACGACGACGACGACGACGACGACGACGACGACG 1174309
QY 135 erPheProSerProSerArgGlyGluProAenAenAenMetSerSerThrPhePhePro 155
DB 1174308 CAGTCCCAACCCCACTGCTGCTCCGCTCATGACGACGACGACGACGACGACGACGACGACG 1174259
QY 155 heLeuArgAnGlyGlyIleProSerSerLeuProSerLeuArgIleSerAenSerCyAs 175
DB 1174258 -----CCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1174216
QY 175 roValThrProProValSerSerProThrSerIleAenProLysProLeuProAenTrpG 195
DB 1174215 CACGACACCCACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1174169
QY 195 luSerIleAlaLysGlnSerMetAlaIleAlaLysGlnSerMetAlaSerPheAenTyrP 215
DB 1174168 -----CCGGCTTCTGACGACGACGACGACGACGACGACGACGACGACGACG 1174153
QY 215 roPheTyrAlaValSerAlaProAlaSerProThrIleArgIleGlnPheHisThrLeuA 235
DB 1174152 CACGACCTGGACACCACTACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 1174096
QY 235 laThrIleProGluCyAspArgGlySerSerSerThrValAspSerGlyHisTrpIleS 255
DB 1174095 CGACCTGCCACCT--ACCCCTTCCACACGACGACGACGACGACGACGACGACGACGACG 1174046
QY 255 erPheGlnLysPheAlaGlnGlnGlnProPheSer-----AlaSerMetValProT 272
DB 1174045 -----CCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1174000
QY 272 hrSerProThrPheAenLeuValLysProAlaPro 283
DB 1173999 CCACCCCTCTCCACGCAACACTCCTCCGCA 1173965

RESULT 15
US-10-029-386-20619
; Sequence 20619, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20619
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002070.1 SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2

```

```

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
US-10-029-386-20619

Alignment Scores:
Pred. No.: 0.000548 Length: 506
Score: 145.00 Matches: 59
Percent Similarity: 41.67% Conservative: 21
Best Local Similarity: 30.73% Mismatches: 81
Query Match: 8.10% Indels: 31
DB: 13 Gaps: 3

US-09-995-938a-7 (1-336) x US-10-029-386-20619 (1-506)
QY 103 SerArgValThrProTyrSerSerGlnAenGlnSerProLeuSerSerAlaPheGlnSer 122
DB 5 TCATCATCATCATCATCATCTCTCATCATCATCATCATCATCATCATCATCATCATCATCA 64
QY 123 ProIleProSerTyrGlnValSerProSerSerSerPheProSerProSerArgGly 142
DB 65 CCATCACCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCTCA 124
QY 143 GluProAenAenAenMetSerSerThrPhePheProPheLeuArgAnGlyIlePro 162
DB 125 CCACCATCATCATCATCATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 184
QY 163 SerSerLeuProSerLeuArgIleSerAenSerCyProValThrProProValSerSer 182
DB 185 TCATCTTCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 244
QY 183 ProThrSerLysAenProLysProLeuProAenTrpGluSerIleAlaLysGlnSerMet 202
DB 245 CCATCATCTTCCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 277
QY 203 AlaIleAlaLysGlnSerMetAlaSerPheAenTyrProPheTyrAlaValSerAlaPro 222
DB 278 -----TCCTCACCAACCATCATCATCATCATCATCATCATCATCATCATCA 307
QY 223 AlaSerProThrHisArgHisGlnPheHisThrLeuAlaThrIle-ProGluCyAsp 241
DB 308 CCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 367
QY 242 GluSerAspSerSerThrValAspSerGlyHisTrpIleSerPheGlnLysPheAlaGln 261
DB 368 CCATCA---TCATCACCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 424
QY 262 GlnGlnProPheSerAlaSerMetValProThrSerProThrPheAenLeuValLysPro 281
DB 425 TCATCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 463
QY 282 AlaProGlnGlnMetSerProAenThrAla 291
DB 464 ---CCATCATCATCATCATCATCATCATCATCATCATCA 490

```

Search completed: December 20, 2003, 07:57:34
Job time : 3656 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 19, 2003, 22:30:12, Search time 4384 Seconds
(without alignments)
3135.409 Million cell updates/sec

Title: US-09-995-938A-7
Perfect score: 1765
Sequence: 1 MTSDDGATSTGAAAAA.....IHDVGMEDLETLGNGARG 336

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US0995938/funat.19122003.142606.4237/app.query.fasta_1.519
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0995938 @CGN 1.1 3508 @runat.19122003.142606.4237 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTPUT -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEF TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: GenEmbl:
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*
15: gb_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vt:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrc:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------------|---------------------|
| 1 | 1758 | 99.6 | 1501 | 8 AY087257 | AY087257 Arabidops |
| 2 | 1755 | 99.4 | 1011 | 8 AY093747 | AY093747 Arabidops |
| 3 | 1755 | 99.4 | 1476 | 8 AY065049 | AY065049 Arabidops |
| 4 | 1664 | 94.3 | 2687 | 8 AF494338 | AF494338 Arabidops |
| 5 | 1664 | 94.3 | 97263 | 8 AC013258 | AC013258 Arabidops |
| 6 | 1519.5 | 86.1 | 1340 | 8 AY065041 | AY065041 Arabidops |
| 7 | 1513.5 | 85.8 | 1257 | 8 AF134217 | AF134217 Arabidops |
| 8 | 1511.5 | 85.6 | 1359 | 8 AY086340 | AY086340 Arabidops |
| 9 | 1432 | 81.1 | 120977 | 8 AC025808 | AC025808 Genomic s |
| 10 | 1060 | 60.1 | 1494 | 8 AF395901 | AF395901 Lycopersi |
| 11 | 1035 | 58.6 | 945 | 8 AF372937 | AF372937 Arabidops |
| 12 | 872.5 | 49.4 | 897 | 6 AX653881 | AX653881 Sequence |
| 13 | 864.5 | 49.0 | 1389 | 8 AY050394 | AY050394 Arabidops |
| 14 | 860.5 | 48.8 | 957 | 8 AY097357 | AY097357 Arabidops |
| 15 | 856 | 48.5 | 14480 | 8 AP004276 | AP004276 Oryza sat |
| 16 | 825 | 46.7 | 831 | 6 AX505690 | AX505690 Sequence |
| 17 | 825 | 46.7 | 831 | 6 BT006310 | BT006310 Arabidops |
| 18 | 825 | 46.7 | 1273 | 8 BT002452 | BT002452 Arabidops |
| 19 | 788 | 44.6 | 84196 | 8 AT7345 | AT7345 Arabidops |
| 20 | 788 | 44.6 | 89469 | 8 AT7183 | AT7183 Arabidops |
| 21 | 721.5 | 40.9 | 198354 | 8 ATAP22 | ATAP22 Arabidops |
| 22 | 721.5 | 40.9 | 198780 | 8 ATCHR1V86 | ATCHR1V86 Arabidops |
| 23 | 642.5 | 36.4 | 411 | 8 AY074829 | AY074829 Arabidops |
| 24 | 585.5 | 33.2 | 978 | 8 AY090331 | AY090331 Arabidops |
| 25 | 585.5 | 33.2 | 1616 | 8 AY050430 | AY050430 Arabidops |
| 26 | 561.5 | 31.8 | 1116 | 6 AX65341 | AX65341 Sequence |
| 27 | 541 | 30.7 | 855 | 6 AX651445 | AX651445 Sequence |
| 28 | 541 | 30.7 | 1217 | 8 AY088379 | AY088379 Arabidops |
| 29 | 541 | 30.7 | 1258 | 8 AK118850 | AK118850 Arabidops |
| 30 | 461.5 | 26.1 | 90425 | 8 F9K20 | F9K20 Arabidops |
| 31 | 416.5 | 23.6 | 15815 | 8 AP003105 | AP003105 Oryza sat |
| 32 | 339.5 | 19.2 | 192110 | 2 AP004070 | AP004070 Oryza sat |
| 33 | 326.5 | 18.5 | 166126 | 2 AP003686 | AP003686 Oryza sat |
| 34 | 313.5 | 17.8 | 176734 | 2 AP003682 | AP003682 Oryza sat |
| 35 | 299 | 16.9 | 11911 | 8 ATFL135 | ATFL135 Arabidops |
| 36 | 299 | 16.9 | 199075 | 8 ATCHR1V49 | ATCHR1V49 Arabidops |
| 37 | 293.5 | 16.6 | 133330 | 2 OSJN00266 | OSJN00266 Arabidops |
| 38 | 293.5 | 16.6 | 148951 | 2 OSJN00241 | OSJN00241 Arabidops |
| 39 | 283.5 | 16.1 | 161277 | 8 AP002972 | AP002972 Oryza sat |
| 40 | 255 | 14.4 | 374 | 8 AY200612 | AY200612 Arabidops |
| 41 | 233 | 13.2 | 163095 | 8 AP002523 | AP002523 Oryza sat |
| 42 | 204.5 | 11.6 | 2070 | 8 BT006482 | BT006482 Arabidops |
| 43 | 204.5 | 11.6 | 2272 | 8 AK117140 | AK117140 Arabidops |
| 44 | 201 | 11.4 | 125008 | 2 AP005546 | AP005546 Oryza sat |
| 45 | 195.5 | 11.1 | 109135 | 8 AP003408 | AP003408 Oryza sat |

RESULT 1

ALIGNMENTS

| LOCUS | AY087257 | 1501 bp | mRNA | linear | PLN 14-APR-2003 |
|------------------------|---|---------------|-------|--------|-----------------|
| DEFINITION | Arabidopsis thaliana clone 33367 mRNA, complete sequence. | | | | |
| ACCESSION | AY087257 | | | | |
| VERSION | AY087257.1 | GI:21405981 | | | |
| KEYWORDS | FLI CDNA. | | | | |
| SOURCE | Arabidopsis thaliana (chale crease) | | | | |
| ORGANISM | Arabidopsis thaliana | | | | |
| REFERENCE | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis. | | | | |
| AUTHORS | 1 (basea 1 to 1501) Haae,B.J., Volkovsky,N., Town,C.D., Troukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L. | | | | |
| TITLE | Full-length messenger RNA sequences greatly improve genome annotation | | | | |
| JOURNAL | Genome Biol. 3 (6), RESEARCH0029 (2002) | | | | |
| MEDLINE | 22088475 | | | | |
| PUBMED | 12093376 | | | | |
| REFERENCE | 2 (bases 1 to 1501) | | | | |
| AUTHORS | Broyer,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K. | | | | |
| TITLE | Full-length cDNA from Arabidopsis thaliana | | | | |
| JOURNAL | unpublished | | | | |
| REFERENCE | 3 (bases 1 to 1501) | | | | |
| AUTHORS | Broyer,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA | | | | |
| COMMENT | This clone sequence is one of 5,000 Cereia full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Wa or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Geneset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly. | | | | |
| FEATURES | Location/Qualifiers | | | | |
| Source | 1..1501 | | | | |
| | /organism="Arabidopsis thaliana" | | | | |
| | /mol_type="mRNA" | | | | |
| | /db_xref="taxon:3702" | | | | |
| | /clone="33367" | | | | |
| CDS | 187..1197 | | | | |
| | /codon_start=1 | | | | |
| | /product="unknown" | | | | |
| | /protein_id="AA064812.1" | | | | |
| | /db_xref="GI:21592862" | | | | |
| | /translation="MTSDGATSTSAAAAAAAAAARPKSRRENNRRRRRRARA AKITGRLAOGDNYLPKRCDDNNEYLKALCEAGVVEDGTTTKCKPLDGLAGTS SRVTPYSQNSQPLSAPFQSPISPVQSPSSSPSPSRGPNKNSSTFFPLRNGG IPSSLPLSLRINSNCVPTPQSPSPSKPLPNNEJAKQSMALAKNSMSPFNFFAA VSAFAPSPTRHQFHPATIPEDCESDSTVDSGHWISFOKAQOQPSASVPTSPFF NLVKAAPQOMSPNTAAFOETQSSBFKENSQVPMBSERLHDVGMDLELTGNGKAA RG" | | | | |
| BASE COUNT | 389 a | 349 c | 312 g | 451 t | |
| ORIGIN | | | | | |
| Alignment Scores: | | | | | |
| Pred. No.: | 5,59e-73 | Length: | 1501 | | |
| Score: | 1758.00 | Matches: | 335 | | |
| Percent Similarity: | 99.70% | Conservative: | 0 | | |
| Best Local Similarity: | 99.70% | Mismatch: | 1 | | |
| Query Match: | 99.60% | Indels: | 0 | | |

[illegible]

| Qy | 221 | AlaPcAlaSerProThrHisArgHisGlnPheHisThrLeuAlaThrIleProGluCys | 240 |
|------------|--|---|-----------------|
| Db | 866 | GCACCTGCTACTCCGACACATCGCACCACTTTCATACCCCGGCTACTATCTGAAATGT | 925 |
| Qy | 241 | AspGluSerAspSerSerThrValAspSerGlyHisThrIleSerPheGlnIlePheAla | 260 |
| Db | 926 | GATGAGTCTGACTTTCACACTGTTGATCTGTGTCACTTGAAATGACTTTCAGAACTTGGCA | 985 |
| Qy | 261 | GlnGlnGlnPProPheSerAlaSerMetValProThrSerProThrPheAsnLeuValIys | 280 |
| Db | 986 | CACACAGACCATCTCTCGCCCTATGTGTGGCAACCTCTCTACTCTTCATCTGTGAAA | 1045 |
| Qy | 281 | ProAlaPProGlnGlnMetSerProAsnThrAlaAlaPheGlnGlnIleGlyGlnSerSer | 300 |
| Db | 1046 | CCTGGGCTCCAGCAGATCTCTCCAAATCTCTGCTTCCAAAGATTGGTCAAAGCTCT | 1105 |
| Qy | 301 | GluPheIysPheGlnIleAsnSerGlnValIleAspProTrpGlnGlyGlyIleHisAspVal | 320 |
| Db | 1106 | GAGTTTAATATTGAGAAATGACCAATTAACTCTGGGAAGAGAGAGATATCATGATGTG | 1165 |
| Qy | 321 | GlyMetGluAspLeuGlnLeuThrLeuGlyAsnGlyIleValAlaArgGly | 336 |
| Db | 1166 | GGTATGGAGGATCTTTGACCTTACACTTGGAAATGGGAAGCTCTCGT | 1213 |
| RESULT 4 | | | |
| LOCUS | AF94338 | 2687 bp | DNA linear |
| DEFINITION | Arabidopsis thaliana BZR1 protein (BZR1) gene, complete cds. | | PLN 20-MAY-2002 |
| ACCESSION | AF94338 | | |
| VERSION | AF94338.1 | GI:20270970 | |
| KEYWORDS | | | |
| SOURCE | Arabidopsis thaliana (chale crese) | | |
| ORGANISM | Arabidopsis thaliana | | |
| REFERENCE | Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; auticocylodons; core eudicots; rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis. 1 (basea 1 to 2687) | | |
| AUTHORS | Wang,Z.-Y., Nkano,T., Gendron,J., He,J., Chen,M., Vafeados,D., Yang,Y., Fujiohda,S., Yoshida,S., Asami,T. and Chory,J. | | |
| TITLE | Nuclear-localized BZR1 mediates brassinosteroid-induced growth and feedback suppression of brassinosteroid biosynthesis | | |
| JOURNAL | Dev. Cell 2 (4), 505-513 (2002) | | |
| MEDLINE | 21966115 | | |
| PUBMED | 11970900 | | |
| REFERENCE | 2 (basea 1 to 2687) | | |
| AUTHORS | Wang,Z.-Y. and Chory,J. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (22-MAY-2002) Plant Biology, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Rd., La Jolla, CA 92037, USA | | |
| FEATURES | | | |
| source | Location/Qualifiers | | |
| | 1..2687 | | |
| | /organism="Arabidopsis thaliana" | | |
| | /mol_type="genomic DNA" | | |
| | /db_xref="taxon:3702" | | |
| | /chromosome="1" | | |
| | /map="between nsalll and ACATPaee" | | |
| | /note="ecotype: Columbia" | | |
| gene | 1..2687 | | |
| | /gene="BZR1" | | |
| promoter | 1..880 | | |
| | /gene="BZR1" | | |
| mrna | join(881..1369,1714..2687) | | |
| | /gene="BZR1" | | |
| | /product="BZR1 protein" | | |
| | join(1103..1369,1714..2457) | | |
| | /gene="BZR1" | | |
| | /codon_start=1 | | |
| | /product="BZR1 protein" | | |
| | /protein_id="AA18490.1" | | |
| | /db_xref="GI:20270971" | | |
| | /translation="WTSGATSTGAAAAAARRKPSREEDNNRRERRRRAYVA | | |
| | KITGALAQGDYNIWKCDNNEVTKALCVAGWVEEDGTTYKGGCKPLGEIAGTS | | |
| CDS | | | |

[illegible]

QY 146 AaAanMetSerSerThrphephepProphelaAaAaGlyGlyIleProSerSerIleu 165
 DB 1882 AACAAATATCTCTTACATTTCTCCCTTCTCAGAAATGGGCAATCTTCTTCTCT 1941
 QY 166 ProSerIleuAaGlyIleSerAaSerCybProValThrProProValSerSerProThrSer 185
 DB 1942 CCTTCCCTCAGAAATCTCAAAAGTTTCCAGTTACCCGAGTCTCATCGCCGACTTCT 2001
 QY 186 IysAaAaProLysProLeuProAaAaThrPGLuSerIleAlaIleGlnSerMetAlaIleAla 205
 DB 2002 AAGAAACCCGAAACCGTGGCTAACTGGGAATCATGCTTACGCAATCATGCGCATTTGCT 2061
 QY 206 IysGlnSerMetAlaSerPheAaAaThrProPheTyrAlaValSerAlaProAlaSerPro 225
 DB 2062 AACAAATCAATGGCGCTTTTAAATATCTTCTTAAGCGGTTTCTCACCCTGCTACTCTCG 2121
 QY 226 ThrHisArgHisGlnPheHisSerThrlaAaIleThrIleProGluCybAaPGLuSerAaPser 245
 DB 2122 AACATCGCCACCAAGTTTCAATACCCGGCTACTATCTGAAATGTATGATGCTGACTCT 2181
 QY 246 SerThrValAaPserGlyHisThrIleSerPheGlnLysPheAlaGlnGlnGlnProPhe 265
 DB 2182 TCACATGTTGATTTCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2241
 QY 266 SerAlaSerMetValProThrSerProThrPheAaAaLysValLysProAlaProGlnGln 285
 DB 2242 TCTGCTCTATGATGGTGAACCTCTCTACCTTCAATCTTGTGAACCTGCGCTCTGACGAG 2301
 QY 286 MetSerProAaAaThrAlaAlaPheGlnGlnIleGlnSerSerGlnPheLysPheGln 305
 DB 2302 AGTCTCCAAATACCTGCTGCTTCCAAAGATTTGGCAAGCTCTAGTTTAAATTTGAG 2361
 QY 306 AaAaSerGlnValLysProThrPGLuGlyGlyAlaGlyHisAaPValGlyMetGlyAaPLeu 325
 DB 2362 AATAGCAAGTTTAAACCTCGGAGAGAGAGAGATCATGATGATGATGATGATGATGAT 2421
 QY 326 GluLeuThrLeuGlyAaAaGlyLysAlaAaGly 336
 DB 2422 GAGCTTACACTTGGAAATGGAAAGGCTCGTGGT 2454
 RESULT 5
 AC013258/c
 LOCUS AC013258 97263 bp DNA linear PLN 19-JAN-2001
 DEFINITION Arabidopsis thaliana chromosome 1 BAC P9E10 genomic sequence,
 complete sequence.
 AC013258
 AC013258.5 GI:12323880
 HTG.
 SOURCE Arabidopsis thaliana (chale cross)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 97263)
 Lin.X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.J.,
 Wu,D., Magill,R., Romling,C.M., Koo,H., Fujii,C.Y., Utecherback,T.R.,
 Bartscheid,M.E., Bowman,C.L., White,O., Niernann,W.C. and Fraser,C.M.
 Arabidopsis thaliana chromosome 1 BAC P9E10 genomic sequence
 Unpublished
 2 (bases 1 to 97263)
 Lin.X. and Kaul,S.
 Direct Substitution
 Submitted (05-NOV-1999) The Institute for Genomic Research, 9712
 Medical Center Dr. Rockville, MD 20850, USA, xlin@tigr.org
 3 (bases 1 to 97263)
 Town,C.D. and Kaul,S.
 Direct Substitution
 Submitted (19-JAN-2001) The Institute for Genomic Research, 9712
 Medical Center Dr. Rockville, MD 20850, USA, cdtown@tigr.org
 On Jan 19, 2001 this sequence version replaced gi:12280837.
 COMMENT Address all correspondence to:at@tigr.org

FEATURES

source

BAC clone P9E10 is from Arabidopsis thaliana chromosome 1.
 The orientation of the sequence is from SP6 to T7 end of the BAC
 clone.
 Genes were identified by a combination of several methods: Gene
 prediction programs including GenScan+ (Chris Burge,
<http://CCR-081.mit.edu/GENSCAN.html>), GenemarkHM (Mark Borodovsky,
<http://genemark.biology.gatech.edu/genemark/>), Glimmer4 (a variant
 of Glimmer3, see Mihaila Petrea,
<http://www.tigr.org/softlab/glimmer4.htm/glimmer4.html>), and
 GeneSplicer (Mihaila Petrea and Steven Salzberg, contact
mperle@tigr.org), searches of the complete sequence against a
 peptide database and the plant EST database at TIGR
<http://www.tigr.org/cdb/tgi.shtml>). Annotated genes are named to
 indicate the level of evidence for their annotation. Genes with
 similarity to other proteins are named after the database hits.
 Genes without significant peptide similarity but with EST
 similarity are named as unknown proteins. Genes without protein
 or EST similarity, that are predicted by more than two gene
 prediction programs over most of their length are annotated as
 hypothetical proteins. Genes encoding tRNAs are predicted by
 tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).
 Simple repeats are identified by RepeatMasker (Arian Smit,
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).
 Location/Qualifiers

1..97263
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="P9E10"
 14..1286
 /gene="P9E10.1"
 <14..1286
 /gene="P9E10.1"
 14..1201
 /gene="P9E10.1"
 /codon_start=1
 /product="unknown protein; 14-1201"
 /protein_id="AAG51915.1"
 /db_xref="GI:12323889"
 /translation="MENCKPQDGTGLOVTLLEVHVVRVRYLTATDLSGLTFTPE
 NRTYGVSPTRPLVFLKRLPLFETAGSLDRSMKIRSEBCGHNLSRSTYF
 DSEPAKAYGFTSEBIIHVLGDMNKCVRSKKYOMESPVALQAKGLVLN
 OEKRVYVNSTOHYVITGFRLLFPALADISTFLSHRESTKTTHOKLEKTPILA
 SPREKLIVMGLDGVAATYKSLPISKAFNNMLWSSPVFPFLIFFGAHHFSSKKE
 SLTAGPDDPFSTTSSSTTTAONSSAFSESTRNDMDLRREYVSSRYPPGA
 TGAHRSVGSNDSSRAPIVETTYRTTAQMKRGGSGDSCGKGRRESLFGNNKALD
 DES"
 1456..3294
 /gene="P9E10.2"
 /note="similar to cytochrome P450 GB:AAA17732
 (Catharanthus roseus); contains Pfam profile: PF00067
 Cytochrome P450"
 join(<1456..1944,2099..2343,2436..2799,2875..3294)
 /gene="P9E10.2"
 join(1456..1944,2099..2343,2436..2799,2875..3294)
 /gene="P9E10.2"
 /codon_start=1
 /product="putative cytochrome P450; 1456-3294"
 /protein_id="AAG51924.1"
 /db_xref="GI:12323898"
 /translation="NAVFPILVLFVFPFLVFRPTYSNIWPMRIQSHFKKOSVTGPSYR
 IFSGSGEVSRLTAABKSPIDSGNPHFVRVAVAHYHMSRVYKTFYAFGSKPV
 VATSPRLIREALTGGSFDRIGHNPLISLVAQGLRGIDGMARHRRARAKAFME
 KLRKVPQWVSTPMLEKMDMRNGEIELEEVHEMNLSAEMSTRPFGSVBEG
 KIFELQERMRPLFLVRSVITPGRRFPSTNREIMRIEKOIRVSLILENNKTA
 VEGSGTLQAFSPYTNQGOEKLGIEIVTECKTFPAAKETTNLMTFVLILAM
 NOEMQNIAREEYICVIGOTGPLILQOLKLSMILETLRLYPAMTLNRLDTLKA
 KLGDIDIPAGTOLSVVMMHDKETWGDSEFNREDPKQALAVFGLGPRT
 CVGQDLAVNEAKTVLATILIKYISPLSPSYAAAPVLVTLQONGHILFTRTSS"
 4711..6345
 /gene="P9E10.3"
 gene


```
mRNA
CDS
    join(<4711..4845,5272..6345)
    /gene="P9E10.3"
    join(4711..4845,5272..6345)
    /gene="P9E10.3"
    /product="unknown protein; 4711-6345"
    /protein_id="AAG51930.1"
    /db_xref="GI:12333904"
    /translation="MAVRKEKVPQPPRECGLIAVLVGLFGICVCTTILIPDFTNFRSS
    KVASCSCEPVRMFAEPAIISKEKGEIARKQVSDLEKVRLEKQEVIAKGPFTV
    TGQVPTVAPEBSANPLAKLEKVAENKEIIVLANNVPMLEVOIAVKVAGVIG
    NYLVPLDLSCEKGEVAYYKRDNDADIVDYSMDGSDVSGLEKRLVRELDLQ
    GYVLSDVDIVFLQNPGRHLXRDSDVMSGDHNNMTAYGPNVDPDPTMRSTVYT
    NRIWFRNSGPPYLRTPLSIELDRLVDTLSKSGNDQAFNQLPTPSHGTYGLYA
    SKAVMDVTEPMNSKVLPTVKRDEMKKLKPVIIHNNYHSDKLERMQAAVEFYNGKQ
    DALDFRDS"
    7482..9019
    /gene="P9E10.4"
    join(<7482..7655,7907..9019)
    /gene="P9E10.4"
    join(7482..7655,7907..9019)
    /gene="P9E10.4"
    /product="unknown protein; 7482-9019"
    /protein_id="AAG51917.1"
    /db_xref="GI:12333891"
    /translation="MAGRDRIQQLKRSRIALIPVILIGCVCSVLFPNGFNGSSS
    LIAERIKSTSTGDLASCESESRVMLKDSFIISVKNALRKQVRELTEKRLAE
    QETENARKQVLVGESEIKAGPFGVQLRTNPTVPDESVPRLAKLEKVAKEII
    VLVANSNVKPMLEIQLASVKRGLIONLIALDSDMSPECSEKVPYKDDPDAVM
    VGGSGNHVAVSGIKRFLVRLREPLQVGLSDVDIVLQNPESHLDSDVESDCH
    DNNTAYGPNVPEPSMGARVATNTKIVFNSGFLRLRTPLSITLLEVDYADTLSK
    BANOAVENQLFTPSHGTYGLHAKRVVMDVMEFMSKVLFTVRNQLKLRPI
    VLVNTHDKLKLRHMAVEPYNGKDALDSFPDSD"
    9598..12387
    /gene="P9E10.5"
    join(<9598..10007,10087..10712,10801..11195,11275..11389,
    11467..11516,11613..11678,11776..11853,11944..12120,
    12221..12387)
    /gene="P9E10.5"
    join(9598..10007,10087..10712,10801..11195,11275..11389,
    11467..11516,11613..11678,11776..11853,11944..12120,
    12221..12387)
    /gene="P9E10.5"
    /product="unknown protein; 9598-12259"
    /protein_id="AAG51921.1"
    /db_xref="GI:12333895"
    /translation="MOTLPSSEVTVLGSNAPPVLRSPGDDVDIDFGVPGCPKRR
    SKVTSNEVTHSFSSEALRRDIVDVGDLIPDEKVFGEIDTSVRRRTTDDFDD
    IFRVNSSSSLPGRSILSPAKHPESSGCTSPQGVFLKAKATEIPTPMLAATSLNKK
    ETVSSPILRTSSKADVSTAKSYDDCDPQVFLKGRQFHSIYKPNKGVPU
    IMGSSRLSSMKABETTPVPLSDYRKTSVYKLGKNEGGKGLGLKQVKTSLKR
    PGVOTKEKTEETDLKSOAPGVSKABEAVKPLDSVESQAPSGVSKAEATTVKRL
    HSIPIHEDERODEKIVSEREVRKSGKAKAKRSTFESRTKTKKSQGTSLDSSPLD
    KSSPASSAAPEVGKGVKGVSDPVKIFKSGAVGAGSGISFMRKKEPKDI
    IHDSNAKETVNIIPDOQKSTPDI PANNRDOQKOSTOKXDSRESNMYAKPGTVE
    EROESTTHTTSEIDEPVNPVEDIDTIDENMEANDAEIKINIDKIRKMSG
    KSGNIRSLSTLQYILMSGSGWVPLMDMIEGNAVKSGVQRALLIHPKLOQKGS
    ANQYMAKRFELQEMHFNLTGPV"
    complement(14275..14526)
    /note="DRI Non-autonomous DNA transposon [a
    consensus]"
    14940..15720
    /gene="P9E10.6"
    /note="similar to putative DNA-3-methyladenine glycosylase
    I (DAG I) GB-P05100 [Baccharifolia coli]"
    join(<14940..15100,15179..15320,15400..15720)
    /gene="P9E10.6"
    join(14940..15100,15179..15320,15400..15720)
    /gene="P9E10.6"
    /codon_start=1
    /product="putative DNA-3-methyladenine glycosylase I;"
```

```
14940-15720"
    /protein_id="AAG51925.1"
    /db_xref="GI:12333899"
    /translation="MVIDQMDPIVLPHDEMGVVRDQKLLFELVFSQALAEFSWP
    SIIRRDPRKLPDEPDPSALIAQETEKRLMSLVNGCLISEOKLRAIVENKSVLK
    KOEFGSNKCMFVNHKPLRNGRYRGVQVRSKPAVEIISKDMWRGRCVGMVY
    SFIQNSGIVNDHLIACPRYQKCVETEBETSHETKLDHSLPV"
    complement(16304..17716)
    /gene="P9E10.7"
    complement(join(<16304..17038,17392..17716))
    /gene="P9E10.7"
    complement(join(16304..17038,17392..17716))
    /gene="P9E10.7"
    /codon_start=1
    /product="unknown protein; 17658-16304"
    /protein_id="AAG51929.1"
    /db_xref="GI:12333903"
    /translation="MNSDGAHSTSGAAAAAARRRKPSREERNRRRRRAVA
    AKIYTGARAQGVNLRPKHCDNNEVLKALCYEAGVNVEDGTTTKPLFELAGTSSRV
    TPTSQMOSPLSSAFQSPITFQVSPSSSPSSRQBPNNMNSTEFPFLNGGIS
    SLPSLRISNCPVTPVPSPTSKNPKPLPMWESIAKSMALAKOSMSFNYPYVSA
    PASPTRHQHPFTPATIIECDSDSDSTVSGSWISFOKFAQQQPSASVPTFNLV
    KPAQQMSPTTAFOEIGOSSEPFEXSQQVKPMWGERIHVDVGMEDLELTNGKARG"
    complement(18764..18835)
```

gene

Alignment Scores:

| Score: | 1,27e-66 | Length: | 97263 |
|------------------------|----------|---------------|-------|
| Percent Similarity: | 1664.00 | Matches: | 334 |
| Best Local Similarity: | 74.06% | Conservative: | 0 |
| Query Match: | 94.28% | Indels: | 2 |
| DB: | 8 | Gaps: | 116 |

US-09-995-938a-7 (1-336) x AC013258 (1-97263)

| | | | |
|----|-------|--|-------|
| Qy | 1 | MetThSerAepdGlyAaThrSerThSerAlaAlaAlaAlaAlaAlaAla | 20 |
| Db | 17658 | ATGACTTGATGATGAGTACGTCGACATCGACCTCAGCTGCGCGGCGAGCAGC | 17599 |
| Qy | 21 | AlaArgArgLysPProSeRTPAArgGluArgGluValuenaAgaArgGluArgArg | 40 |
| Db | 17598 | GCGAGAGGAAAGCCGTCGTGAGAGAAAGGAGAAATATCGAGAGAGAAAGCAGAA | 17539 |
| Qy | 41 | ArgAlaValAlaAlaLysIleTyThGlyLeuArgLagInglYasPTyrAsnLeuPro | 60 |
| Db | 17538 | AGAGCTGTACCTGGAAGATATACACTGGGCTTACAGCTCAAGGTGATTAATTTGCT | 17479 |
| Qy | 61 | LyeHiScYabphenaAenGluValLeuLysAlaLeuCybValGluAlaGlyTrpValAla | 80 |
| Db | 17478 | AAACATTGTGATTAATATGAAGTCTTAAAGCTCTTGTGTGAAGCTGTGGGTGTT | 17419 |
| Qy | 81 | GluGluAspGlyThrThrTyArgLys- | 89 |
| Db | 17418 | GAGAGAGATGATCTACTTATTCGAA-GGTGAAGACTTTCTCCATTTTTCACAGTCTGA | 17360 |
| Qy | 89 | ----- | 89 |
| Db | 17359 | GCTTGTATTATATGTTTGTGAATGTTGAATCGAATTCGATTTCAATTGTGCTTA | 17300 |
| Qy | 89 | ----- | 89 |
| Db | 17299 | AATGGTGTGAATCTGAGAAATTGAGGATTTTCTCAAGAGTAATTGATCATCAGAAAC | 17240 |
| Qy | 89 | ----- | 89 |
| Db | 17239 | TATGATGATCTGATTTCTCAAGAGTAATTATGAGTTTCTTCTTAATTTAGAGTTA | 17180 |
| Qy | 89 | ----- | 89 |
| Db | 17179 | TTATGTGATGCTAAAGTCTTAATCTTTATGTATGATCTGTGCTCAAGTCAATTCAT | 17120 |
| Qy | 89 | ----- | 89 |

| | | | | |
|------------|---|--|--|-------|
| Df | | 17119 | GGTGTTCTTTCCTTACTCGTAATGAATTGATGGTTATGTTTGCTTTTG | 17086 |
| Oy | 90 | -----GlyCybVybProLeuPproglYguilLealagIYhrSerSeraArgval | 105 | |
| Df | 17059 | TTCAGATACAGGGGATGCMAAGCCTTTTACCTGGTAGATACCTGGGACTTATCATCCGAGTA | 17000 | |
| Oy | 106 | ThirProYrSaSerSeGiNaSngLiSerProLeuSerSezAlaPhegInserProIllePro | 125 | |
| Df | 16999 | ACTCCATATTCATCAACGAAACCAGGCCCTTTTCATCACCTTTCMAAATCCCATTCCCA | 16940 | |
| Oy | 126 | SerTYrGiNaISerProSerSerSerSerPheProSerProSeraArgIYguIProAn | 145 | |
| Df | 16939 | TCTTACCAAGTTAGCCGCTCTTTTCATCATCTCCGAGTCTTTCCGGGTGAACCAAT | 16880 | |
| Oy | 146 | AmbAnMeCSeSerThrPhePheProPheLeuArgAngLiYgiYlIeProSerSeLau | 165 | |
| Df | 16879 | AACAACATGATCCTCTACATCTTCCCTTTCCTCAGAATAGTGAGCATTTCTTCTTCTT | 16820 | |
| Oy | 166 | ProSerLeuAggIIesSerAnsSerCybProValThrProProValSerSerProThrSer | 185 | |
| Df | 16819 | CCTTCCCTCGAAMCTCAAACAGTTGTCAGATTACCCACCGGCTTCATCGCCGACTTCT | 16760 | |
| Oy | 186 | LYbaSnProlYbSProleuProAnTrigIuseRtleAlaLyvGinsereMetaIalaIlea | 205 | |
| Df | 16759 | AAGAACCAGAACCGTTGCTTACCTGGGAATCTATCCCTTAAGCAATCATAGCCATTGCT | 16700 | |
| Oy | 206 | LYaGInsereMetaIsaPheAenTYrProPheTYrAlaValISeralAProlaserPro | 225 | |
| Df | 16699 | AAACATCAATAGGGGCTCTTTTAATTATTCCTTTCATCGGGTTCTGCACCTGTAGTCCG | 16640 | |
| Oy | 226 | ThriSarighIsgInPheHisThrLeuAlaThrTlleProGLCybApGiusaRaePseR | 245 | |
| Df | 16639 | ACACATCCCCCAGTTTCATACCCCGGCTACTATACCTCAATGTGATGCTGACTCT | 16580 | |
| Oy | 246 | SeArhVaIaPaSaSeRgiYHiaSTrioLeserPheGlnLYspheLaGnIngInIpncOpe | 265 | |
| Df | 16579 | TCCACTGTGATCTGTGATTCATGGATAAGCTTTCAGAAGTTTGCAACAACAACGCCATTC | 16520 | |
| Oy | 266 | SeRaIsaSerMetValProThrSerProThrPheAmbLeuValLYsProlaAProgInGln | 285 | |
| Df | 16519 | TCTGGCTCTATAGGGCCAACTCTCTACCTCAATCTTGTAACCTCGGCTCAGAG | 16460 | |
| Oy | 286 | MeSeSerProAnThrrAlaAlaPheGlnGuIileGIYgInseRseRguPheLYspheGU | 305 | |
| Df | 16459 | ATGCTCCAAATACTGCTGCTCTCCAAGAGATGGTCAAAAGCTCTGAGTTTAATTGGAG | 16400 | |
| Oy | 306 | ASnsErGInVALysPRotrNGInGILuaArgLIehIsaSPoValGImetGUsaBleu | 325 | |
| Df | 16399 | AATAGCCAAAGTTAAACCTCGGGAAGAGAGAGATCATGATGGETATGAGAGATCTT | 16340 | |
| Oy | 326 | GlueLeuthrLeuGIYaengLYALaARGLIY 336 | | |
| Df | 16339 | GAGCTTACACTTGGAAATGGGAAGCTCTGTGT 16307 | | |
| RESULT_6 | | | | |
| AY065041 | | | | |
| LOCUS | | 1340 bp | mRNA linear PLN 07-JAN-2002 | |
| DEFINITION | Arabidopsis thaliana Atlg193305/F18014_4 mRNA, complete cde. | | | |
| ACCESSION | AY065041 | | | |
| VERSION | AY065041.1 | GI:18086445 | | |
| KEYWORDS | FLI CDNA. | | | |
| SOURCE | Arabidopsis thaliana (chale crees) | | | |
| ORGANISM | Arabidopsis thaliana | | | |
| | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | | |
| | Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; | | | |
| | Rosids; euroside II; Brassicales; Brassicaceae; Arabidopsids. | | | |
| | 1 (bases 1 to 1340) | | | |
| REFERENCE | | | | |
| AUTHORS | Chouk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banb J., Bayer S.L., Carnunli P., Chang E., Dale J.M., Goldsmith A.D., Hoyewitzak I.Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J., Quech H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., | | | |

Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1340)

Chen, R., Chen, H., Kim, C.J., Koeseema, E., Meyers, M.C., Ban, J., Bowser, L., Carninci, P., Chang, E., Dale, J.W., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.W., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Direct Submission
Submitted (30-NOV-2001) Salk Institute Genomic Analysis Laboratory (Signal), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RALF cDNAs (RALF cDNA : RIKEN Arabidopsis Full-Length cDNATM) : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PECC (SSP) Consortium members carried out the sequencing and annotation of the RALF cDNAs: Chen, R., Chen, H., Kim, C.J., Koeseema, E., Meyers, M.C., Shinn, P., Ban, J., Bowser, L., Dale, J.W., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.W., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Chen, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers

1. 1340

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="1"

/clone="RALF05-03-B11(R21175)"

/note="ecotype: Columbia"

1. 137

138. 1145

/note="unknown protein"

/codon_start=1

/product="Atlg19350/F18014_4"

/protein_id="AL57677.1"

/db_xref="GI:18086446"

/translation="MTSDGATSTSAAAAAAATRRKPSGRRENNRRPRRRRAVA
KRTGARAQGNVTLPRKCDNNELVLAALSEAGVIVDEDTTKRGKRLPGDAGSS
RAIPYSHNOSPLSTFDSPILSYOVSPSSSPSPSRGDPNNISTIFPLNGGIP
SSTPLPRTINSAPAVTPVSPSPSRNRPPLPTWMSFTKQSWMAKOSMTSLNPFYAV
SAAASPETHROFAPATPIECEDSDGSIYDSGMIISFOKFAOQOPASASVPSPTFN
LVHPAPQQLSPNTAAIOEIGOSSEPFKFNSTQVPEGGERIHDVAMEDLETLNGKXAH
S"

3' UTR 1146. 1340

BASE COUNT 357 a 315 c 284 g 384 t

ORIGIN

Alignment Scores:

Pred. No.: 5,01e-62 length: 1340

Score: 1519.50 Matches: 297

Percent Similarity: 91.37% Conservative: 10

Best Local Similarity: 86.39% Mismatches: 24

Query Match: 86.09% Indels: 5

DB: 8 Gaps: 4

US-09-995-938A-7 (1-336) x AY05041 (1-1340)

| Db | 352 | TCATCTTCCGAGCAAC | CTCTTACTCTTCCCATTA | CCAAAGCTCTTCTTCCACTTTT | 411 |
|------------|--|---|--------------------|------------------------|-----------------|
| Qy | 121 | GlnSerProIleProSerTyGlnValSerProSerSerSerPheProSerProSer | 140 | | |
| Db | 412 | GATACCCCCCATCTTATCTTACCAAGCAAGCTCTTCTCTTCTTCAATCCGAGCTCTTCT | 471 | | |
| Qy | 141 | Arg---GlyGluProAsnAsnAsnMetSerSerThrPhePheProPheLeuArgAsnGly | 159 | | |
| Db | 472 | CAAGTTGGTGGATGCCAACAT-----ATCTCACAAATCTTCCCTTCTCTCGAAGATGCT | 525 | | |
| Qy | 160 | GlyIleProSerSerSerLeuProSerLeuArgIleSerAsnSerCysProValThrProPro | 179 | | |
| Db | 526 | GGTATTCCTTATCTGCTCTTCCACTTAAGATCTCAACAGAGCTCTCTGTCACTCCACCA | 585 | | |
| Qy | 180 | ValSerSerProThrSerLeuAsnProIleProLeuProAsnTyrGlnSerIleAlaLys | 199 | | |
| Db | 586 | GTGTCAATCCCCCAATCTTGAAGAACCCCAACCAATGGCTGCTACTGGGAATCTTTACCAA | 645 | | |
| Qy | 200 | GlnSerMetAlaIle---AlaValGlnSerMetAlaSerPheAsnTyrProPheTyrAla | 218 | | |
| Db | 646 | CAATCATCATGTCCATGAGCTGCTCTTAACAGTCAATAGCATCTTTGAACCTACCCGTTATATCG | 705 | | |
| Qy | 219 | ValSerAlaProAlaSerProThrHisArgHisGlnPheHisThrLeuAlaThrLeuPro | 238 | | |
| Db | 706 | GTGTGTGACCCCTGCACAGCTCACTATCATCATGCCAGTTCATCATGCTCCGGCTACTATACCT | 765 | | |
| Qy | 229 | GluCysAspGlnSerAspSerSerThrValAspSerGlyHisTyrIleSerPheGlnLys | 258 | | |
| Db | 766 | GAATGTGATGAGTCTGCAGCTCTTCCACTGTGATTTCTGGTCATGGATAGCTTTCAAAAG | 825 | | |
| Qy | 259 | PheAlaGlnGlnGlnProPheSerAlaSerMetValProThrSerProThrPheAsnLeu | 278 | | |
| Db | 826 | TTTGACACAACAACACCAATCTCTGCTCTATGAGTGCACCACTCCGCTACCTCAATCTC | 885 | | |
| Qy | 279 | ValLysProAlaProGlnGlnMetSerProAsnThrAlaAlaPheGlnGluIleGlyGln | 298 | | |
| Db | 886 | GTGAACCTGCACCAACAGCAATGTGTTCCAAACACAGCAGCATCAAGATGTGTCAA | 945 | | |
| Qy | 299 | SerSerGluPheLeuPheGlnLysSerGlnValLysProTyrGlnGlyGluArgIleHis | 318 | | |
| Db | 946 | AGCTCCGAGTTTAAGTTTGAGAACGCCAAGTTAAGCCATGGGAAGGGAGAGATCAT | 1005 | | |
| Qy | 319 | AspValGlyMetGluAspLeuGluLeuThrLeuGlnGlyAsnGlyLysAla | 334 | | |
| Db | 1006 | GATGGCGCTATGGAGGATTTAGAGCTCAAGCTTGGAATGGTTAAGCT | 1053 | | |
| RESULT 8 | | | | | |
| LOCUS | AY086340 | 1359 bp | mRNA | linear | PLN 14-APR-2003 |
| DEFINITION | Arabidopsis thaliana clone 1359 bp mRNA, complete sequence. | | | | |
| ACCESSION | AY086340 | | | | |
| VERSION | AY086340.1 | GI:21405050 | | | |
| KEYWORDS | FLI CDNA. | | | | |
| SOURCE | Arabidopsis thaliana (thale cress) | | | | |
| ORGANISM | Arabidopsis thaliana | | | | |
| REFERENCE | 1 (bases 1 to 1359) | | | | |
| AUTHORS | Haaß, B., J., Volkovskiy, N., Town, C. D., Troukhan, M., Alexandrov, N., Feldmann, K. A., Flavell, R. B., White, O., and Salzberg, S. L. | | | | |
| TITLE | Full-length messenger RNA sequences greatly improve genome annotation | | | | |
| JOURNAL | Genome Biol. 3 (6), RESEARCH0029 (2002) | | | | |
| MEDLINE | 22088475 | | | | |
| PUBMED | 12093376 | | | | |
| REFERENCE | 2 (bases 1 to 1359) | | | | |
| AUTHORS | Brayer, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K. | | | | |
| TITLE | Full-length cDNA from Arabidopsis thaliana | | | | |
| JOURNAL | Unpublished | | | | |
| REFERENCE | 3 (bases 1 to 1359) | | | | |

AUTHORS Brover V., Tikhonhan M., Alexandrov N., Lu Y.-P., Flavell R. and Feldmann K.

TITLE Direct Submission

JOURNAL Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road, Malibu, CA 90265, USA

COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to RIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and applied exons; five percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these CDNA sequences are derived from the Wa or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Geneset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

| | |
|--------|---|
| source | location/Qualifiers |
| CDS | <pre> 1..1359 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="24157" 112..1119 /codon_start=1 /product="unknown" /protein_id="AA064408.1" /db_xref="GI:21592457" /translaton="MTSDGATVTSAAAAAAMATERKPSMRERENRRRRRRRAVA KIYGLPAQGNYNLPKHDDNEVLALCSBAGAVVEEDTTRKGKPLPGDAGSS RATYSSNNOSLSTFSPISLYOVSSKSLPTPESTTKOSMAAKOSMILNPKYX SLLPLINSAPVPTPPSPSPRSKRSLPTPESTTKOSMAAKOSMILNPKYX SAPSPHHRGFAPAHTIPCEDBSSSTVDSCMHLSFOFAAOOPFSASNVPSPTFR LVKAPPOOLSPNTAIIOEIRIQSSSEKFENSOVKPWEGERIHDVAMELELTGNGKK S"</pre> |

BASE COUNT 356 a 317 c 285 g 401 t

ALIGNIN

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. NO.: | 1..19e-61 | Length: | 1359 |
| Score: | 1511.50 | Matches: | 296 |
| Percent Similarity: | 91.07% | Conservative: | 10 |
| Best Local Similarity: | 88.10% | Mismatches: | 25 |
| Query Match: | 85.64% | Indels: | 5 |
| DB: | 8 | Gaps: | 4 |

US-09-995-938A-7 (1-336) * AY086340 (1-1359)

Oy 1 Methtsrarabpglyalathserhsrseralelaialaiaialaiaialaia 20

Db 112 ATGCACGTCTGACGGAGCAACGTCAACGTCA--GCTGACGCTGCACAGCGAGCGC 168

Oy 21 Alaaargtqlvprosetrtpraggluarvgluuaenaenaraargarggluarvgarg 40

Db 169 ACAGGAGGAACAACGTCCTCGTAGAGAGAGGAGGAGAACATCGGAGAAGAGCGCGGAGA 228

Oy 41 ArgAlaValAlaIalaLyIleTyrrThrgLyLeuaArgAlaGlngLyAsPyraenLeupro 60

Db 229 AGACTGTTGGCGGAGAAATTATTACTGCTTTAAGAGCTCAAGGTAACTAACATCTTCCA 288

Oy 61 LyehIsCySaBaPaNaEnGluValLeuLySaIaLeuCyVaIaGluJaIaGlyTPvvaIval 80

Db 289 AAACATTTGTCACAAATGAGAGTCTTAAAGCTCTTTGTTGCAACCTGTTGGGTGTT 348

Oy 81 GluglUasPgJyThrThTyfArghysgJyCySlvsPProLeuProclYglUlleaIagly 100

Db 349 GAAGAAAGACGGAACCTACTATTCCGACAGGAGCACAAAGCCTCTACCTGTGTACATGCGTGGA 408

Oy 101 ThrsErGrArGaIThrPrOtyrSrGrInaEngInSePrOleuSrGrAlaPhe 120

```

Db      409  TCATCTTCTGAGCAACTCTTACTCTTCCATTAACCAAGCTCTTCTTCCACTTTT 468
Qy      121  GlnSerProIleProSerTyGlnValSerProSerSerSerSerPheProSer 140
Db      469  GATAGCCCCCATCTTATCTTACCAAGTCAGTCTTCTCTTCTTCACTCCGAGTCTTCT 528
Qy      141  Arg---GlyGluProAsnAsnAsnMetSerSerThrPhePheProPheLeuArgAngly 159
Db      529  CGAAGTTGGTGAATCCACACAAAT-----ATCTCCAAATCTTCCCTTCTTCCAGGATG 582
Qy      160  GlyIleProSerSerLeuProSerLeuArgIleSerAsnSerCysProValThrPro 179
Db      583  GGTATTCCTTATGCTCTTCTTCCACTTGAATTCACCAAGTCTCTCTCTCTCTCCACCA 642
Qy      180  ValSerSerProThrSerTyAsnProIleProIleProIleProIleProIleAllys 199
Db      643  GTGTATCTCCCACTTCTTGAACCTCCAAACCATGCTGCTTGGGAATCTTTACCAAA 702
Qy      200  GlnSerMetAlaIle---AlaIleGlnSerMetAlaSerPheMetTyProPheTyAla 218
Db      703  CAATTCATGTCATGCTGCTGCTTAAACATGCAATGACTTCTTGAACCTACCCGTTTATGCG 762
Qy      219  ValSerAlaProAlaSerProThrHisArgHisGlnPheHisThrLeuAlaThrIlePro 238
Db      763  GTGCTGTCACCTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 822
Qy      239  GluCysAspGluSerAspSerSerThrValAspSerGlyHisIlePheSerPheGlnlys 258
Db      823  GAATGTATGATGATGCTGCTTCTTCACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 882
Qy      259  PheAlaGlnGlnGlnProPheSerAlaSerMetValProThrSerProThrPheAsnLeu 278
Db      883  TTGTCACAAACAAAGCCATCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 942
Qy      279  ValIleProAlaProGlnGlnMetSerProAsnThrAlaIlePheGlnGlnIleGlyGln 298
Db      943  GTGAAACCTGCACACACAAATGCTTCTTCAACACAGACAAATCCAAAGATGCTTCTTCT 1002
Qy      299  SerSerGluPheTyPheGluAsnSerGlnValIleProIleProIleProIleProIleHis 318
Db      1003  AGCTCCGAGTTTAACTTTGAACACAGCCAACTTAAAGCATGGGAAGGGAAGATTCAT 1062
Qy      319  AspValGlyMetGlyAspLeuGluLeuThrLeuGlyAsnGlyIleVala 334
Db      1063  GATGTGCTATGAGAGATCTAGACTCAGCTGTAATGTAAGCT 1110

RESULT 9
AC025808 120977 bp DNA linear PLN 11-OCT-2000
LOCUS Genomic sequence for Arabidopsis thaliana BAC F18014 from
DEFINITION chromosome 1, complete sequence.
AC025808
VERSION AC025808.8 GI:7636235
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 120977)
AUTHORS Shim, P., Brooks, S., Buehler, E., Chao, O., Johnson-Hopson, C.,
Khan, S., Kim, C., Alatafi, H., Bel, O., Chin, C., Chou, J., Choi, E.,
Conn, L., Conway, A., Gonzales, A., Hansen, N., Howling, B., Koo, T.,
Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S.,
Mukharbeky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J.,
Southwick, A., Thaveri, A., Tortum, M., Vaysberg, M., Yu, G.,
Federici, N. A., Theologis, A. and Ecker, J. R.
Genomic sequence for Arabidopsis thaliana BAC F18014 from
chromosome 1
TITL Unpublished
JOURN 2 (bases 1 to 120977)
REFER Ecker, J. R.
AUTHOR

```

```

TITL Direct Submission
JOURN Submitted (15-MAR-2000) Arabidopsis thaliana Genome Center,
AUTHOR Department of Biology, University of Pennsylvania, 38th Street and
REFERENCE Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
AUTHORS 3 (bases 1 to 120977)
Ecker, J. R.
TITL Direct Submission
JOURN Submitted (22-APR-2000) Arabidopsis thaliana Genome Center,
AUTHOR Department of Biology, University of Pennsylvania, 38th Street and
REFERENCE Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
AUTHORS 4 (bases 1 to 120977)
Chen, R., Shim, P., Brooks, S., Buehler, E., Chao, O.,
Johnson-Hopson, C., Khan, S., Kim, C., Alatafi, H., Bel, O., Chin, C.,
Chou, J., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N.,
Howling, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K.,
Liu, S., Mukharbeky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H.,
Schwartz, J., Southwick, A., Thaveri, A., Tortum, M., Vaysberg, M.,
Yu, G., Davis, R., Federici, N., Theologis, A. and Ecker, J.
TITL Direct Submission
JOURN Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
AUTHOR Department of Biology, University of Pennsylvania, 38th and
REFERENCE Hamilton Walk, Philadelphia, PA 19104-6018, USA
AUTHORS 5 (bases 1 to 120977)
Chao, O., Brooks, S., Buehler, E., Johnson-Hopson, C., Khan, S., Kim, C.,
Shim, P., Alatafi, H., Bel, O., Chin, C., Chou, J., Choi, E., Conn, L.,
Conway, A., Gonzales, A., Hansen, N., Howling, B., Koo, T., Lam, B.,
Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharbeky, N.,
Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A.,
Thaveri, A., Tortum, M., Vaysberg, M., Yu, G., Davis, R.,
Federici, N., Theologis, A. and Ecker, J.
TITL Direct Submission
JOURN Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center,
AUTHOR Department of Biology, University of Pennsylvania, 38th and
REFERENCE Hamilton Walk, Philadelphia, PA 19104-6018, USA
AUTHORS On Apr 22, 2000 this sequence version replaced gi:7543634.
COMMENT Location/Qualifiers
FEATURES
SOURCE
1..120977
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/clones="F18014"
90..2159
/note="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F18014.1"
/protein_id="AAF79419.1"
/db_xref="GI:8778411"
/translation="MISFEVAPVFTCSIVNAYCRSGNVDKAMVFAKETSSGLET
NVVTYNSLNGVAMIGVEGTRVLRIMSEGVSRNVVYTSLLKGYCKGKLMSEAHE
VPELLEKKEKLIVADQHWGVLMDCGCRGQIRDAVRVNDMIEIVRTNTTICNSLNG
YKSGQVLEAQLFSPRNDMSLRPDHTTYTIVDGYCRAGYVDEALKICDQMCQEV
PTVNTYVILKGTSGRIGAFHDVLSLMKMLRGNVADIEISCTLEALFKGDFEAM
KLWENVILARGLTDTITLNVMIISGLCKMKNENKEIILDNNVIFRCKPAVQYQALSH
GYKAGNLKXEAFAVKEIMERKGIPTIEMNTLISGAFYKHLKLVADVLELRAGL
TPVYATGALITGCMNIGMIDKAYATCEMEKEKITLVNVCISRIANSLEPLDKIDEA
CLIDKIVDFDLIPGOSIKPELSEAPRTPETTYTILHICCAAGDINKAPLIDEM
VALAGDCKRAGLDEARLKFSDLSSDRPIDETTYTILHICCAAGDINKAPLIDEM
ALAGLIPNITYVALNLKIGLCKLGNDVDAQRLLHLPGKITLPNATITNTLIDGLVKS
NVAEMVNLKEMIEKGLVRSDDQVDVIREVLLDPEVKLSGTGVIEMNSNELYDR
RVSEAVV"
complement(4647..5702)
/note="putative glycosyltransferase-like protein
emb|CAA42905.1; similar to EST gb|A1998490.1"
/codon_start=1
/evidence=not experimental
/product="F18014.2"
/protein_id="AAF79456.1"
/db_xref="GI:8778448"
/translation="MSQHLLILILSLILHKLPSATTTIOKFEARPOFYNSADCPLI
DSESDVDVAKPIFCRRRAVHVAWTLDAITIRGSVAIVSLVLOHSSCPENIVHFAVA
SASADASLRATISSFPYLDFTVYVENVSSVSRLLSSISRSALDCLNVARSYLADL

```

| | | |
|--|--|--|
| CDS | LPCVRRVVYLDLVDLIAALADLGRDVLAPPEYCNANFTSTPTSPWNP LSTPADRAKAYFNTGVMVITLDRMBEGATYSISEMMAMOKRMRYELGSLPPLLY FAGLIKPVNHRNQNGLGGDNPRGLCRDLHPGVSLIHHWSGKGPWALDAGRPPCLD ALMAPIYDLQTPALDS" | complement(9620..10300) /note="similar to Zn finger protein gi 631991i" /codon_start=1 /evidence=not experimental /product="F18014.3" /protein_id="AA79455.1" /db_xref="GI:8778447" |
| CDS | /translation="MSDVPCSSGNDNNNSNFEPCNCLDLAODPIVTLGHLFCM PCLYKMLHLSQSKDPCVCKAVIEDRLVLYGRKSADPRSKIPGSEVNRPGQ RETQAPVPNHHGRCGHHGRCGFPAPMASAPRGNTLSAARGLPSTLPHPH GFPDAAVMAAASGCPHPSPNPFHGGSHMSYQNRHRCQGDHHLILILIVV VPSLFLS" | join(12523..12577,12686..13374) /note="hypothetical protein" /codon_start=1 /evidence=not experimental /product="F18014.4" /protein_id="AA79420.1" /db_xref="GI:8778412" |
| CDS | /translation="MAIPSTSHLPIFIIATCTISVSGTPTLTNHCSTIWPGLT ANGAQLDGGFALASGSVPTVPSGMSGRPMARPYCNPDASGSGCGTGDGSLKLC AGAGAPATTLAFTIGSSGKKAQVDPDVSLVNGYINQKITPQSGSDCKTACV SVNNAICPELDVTPSGVACACACAFNKPCCYCTGASTPATCPNNYSKIFQA CPSAYADLDSSTPTCNAYEISPCS" complement(14020..14136,14236..14304,14402..14464, 14557..14696,14879..15005,16274..16486) /note="unknown protein; similar to EST gb H76493.1" /codon_start=1 /evidence=not experimental /product="F18014.5" /protein_id="AA79454.1" /db_xref="GI:8778446" | |
| CDS | /translation="MLEAVDSGQVNGGFPQIOLPYGDCSFEELSVLRHKVYVTS NNRTKSVLVGLQGVKKA VGLGGMHLVLTNGIEVLQNALSVLEPPTGNEDDLD PNTQRNSGDMTSEDTLKPXKSLRGQSSRSKSTSGSGSTGTPPEPM KYDLSLEMPALNVRHFNLDVAIPNPEKEQLIDVQHHPFQQQDELDQVYFQA ARMKKACFKGKESRNTDNCIS" join(117994..18489,18568..18698,18814..18847,19006..19099, 19170..19347,19441..19479,19616..19715,19855..19994) /note="hypothetical protein" /codon_start=1 /evidence=not experimental /product="F18014.6" /protein_id="AA79421.1" /db_xref="GI:8778413" | |
| CDS | /translation="MAKTDLAQLDLSGIESDEFNWFPLDITVRIITNRSYTRKVPSS AVYSRFPNKSQJLHSHSESNPKRKRKQKNSFHLPSVEQASNLHQARPLSKAH BSFLKEILSLTKGLSDNDSDSLNKKCDDEVSFTELGVMAAPYEITLSNHL HCDNBESNEORGVQVFNNLVNEIGEVEAFSRRIRYIMRNSCPTQSDHJHNTL VAKSEEGYNLVIDPMEWNAHAKSKSVIYALMYPTLISNOVETSLPFIQALHABAL VALATNRKSLSPFEKELPAMGIRKYATMYKTKPDGTLCIDDLIAGSKSDP KLDNRQIIMSIIDGDSRRPPIGDIILKHTPGSQPARCLEFAREMAAGMTWGNBPL HFQDSRYFLK" | join(22037..22300,22618..22361) /note="unknown protein; similar to ESTs gb A1992723.1, gb T41777.1, db AV440980.1, gb A1997341.1, and db AV440785.1" /codon_start=1 /evidence=not experimental /product="F18014.7" /protein_id="AA79422.1" /db_xref="GI:8778414" |
| CDS | /translation="WSDGATSTSAAMAAAATRRKPSRENNRRERRRAVAA KITYTGRAGQNNYLPKRCNNNEVLKALCEAGVVEEDDTYRKGRPLPGMAGSS KATPYESHQSPSLSTPDPILSYQVSPSSSEFSPSRVIGDPHNISTIPFLANGIP SLPLRISNSAPVTPVSGSPSRNKP.LPTWESFTQMSMAKQSLNLYPPAY SAPASTPHNRQFAPATIPBCDSDDSTVDSGHWISFOKFAOQOPASAMVTPSPFN LYKRAPQQLSPNTAAIOEIGQSEEPKENSQVPMGGERIHDVAMDELTLLNGNAH S" | |
| CDS | complement(123868..24983,25237..25407) /note="unknown protein; similar to EST gb T13858.1, emb Z38050.1, and emb Z38087.1" /codon_start=1 /evidence=not experimental /product="F18014.8" /protein_id="AA79453.1" /db_xref="GI:8778445" | |
| CDS | /translation="MAGRDRCQQLRGSRITAIILIGITGCVAVLPYGFPNSSS LKASEHLSKSQVSSGACSPERVMKSDVPTLSEKNAELKQVRELTEKRLAEQ GDNARKQVLTALGTQIKAGPFGTVSLRNPTLIPDESINPRIAKLIEBAVDREYV ALANAVKAMELVQJLASIKRVGITVYLVAALDVIENLCKENVAAYKRDVDV GKTGNHVASGLKPFVLEFLGVLSDVDIVLQMPESHLYRSDVESDGD NHTAGFNDVPEPAMGARVYHTWIRIWPNGGFTLRPTISIELLDVADLSKX VMDQVPELTPSHPEYTLAASRVDLSEPMKSLVLFKVRNHEKKVAPV VHVNHPDLNMQAVERYVNGQDALDLPFGDS" complement(26100..27917) /note="unknown protein; similar to ESTs gb T42717.1, emb F13886.1, and emb F13885.1" /codon_start=1 /evidence=not experimental /product="F18014.9" /protein_id="AA79452.1" /db_xref="GI:8778444" | |
| CDS | /translation="MADPLNGKSEFFITFLPFTLLPISFLYATESPVIEDSTVYA VASVSTNREAVLHLNLELVQNTLVLNLDKLSLTPLEKKEIEVDDDIGEEKRG KAKASVYKSPFMSERFOGTSAVNSATQINVLPEPDESSSYFALGDSGRVY VLRNGDVIIEFTVDSVPTAMVSYSVFKNSFPVTHQONAVLHRIHESNGD WNSNSVMEHVGKFPVDSADPVTLELVHVRVRLTLATDSGKTLVLENTVYS VPSRPLVFLKQRLFLFESGASLDLSMKIRRETCGLNLSLARTVYFDAERSK | |
| Alignment Scores: | | |
| Pred. No. : 8,276-56 Length: 120977 | | |
| Score: 1432.00 Matches: 297 | | |
| Percent Similarity: 69.468 Conservative: 10 | | |
| Best Local Similarity: 67.194 Mismatches: 24 | | |
| Query Match: 81.138 Indels: 112 | | |
| DB: 81 Gaps: 6 | | |
| US-09-995-938a-7 (1-336) x AC025808 (1-120977) | | |
| QY | 1 MetTtRseAaPbG AAlaThrSerThySerAlaAlaAlaAlaAlaAlaAla | 20 |
| Db | 22037 ATGAGCTGTGAGCGAGCAACGTGACGTCA---GCTGACGCTGACACAGAGGATGCG | 22093 |
| QY | 21 AlaAaGArGLyPProSerTrrPArGLuaArgGLuaAsnAaArgArGLuaTArGArG | 40 |
| Db | 22094 ACGAGGAGAAACCGTCGTGAGAGAGAGAGACAACTCGAGAGAGCGCGGAGAG | 22153 |
| QY | 41 ArGaLaVaLaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla | 60 |
| Db | 22154 AAGCGTGTGGCGGCAAGTTTAACTGCTTAAAGCTCAAGGTAACTCAATCTTCCA | 22213 |
| QY | 61 LysHLSQVAsPAsnAsnGLuValLeuYsAlaLeuCySeVaGLuLaGLyTTPvaLva | 80 |
| Db | 22214 AACAATGTGACAACATAGGTTCTTAAGGCTTTGTTGTAAGCTGGTGGTGGT | 22273 |
| QY | 81 GluGLuaPGLYThrThrTYrArGLySGLY----- | 90 |
| Db | 22274 GAAAGAACGGAACACTACTATGCGAA--GGTCAGTATCAAAAGCATTTTACTTAGATCTG | 22332 |
| QY | 90 ----- | 90 |
| Db | 22333 ATGTTATCTGATGATTAGCTGTGAATCTGAAGATTGGATTGAAATTGTCAAATTG | 22392 |
| QY | 90 ----- | 90 |
| Db | 22393 GGAATTTCTTGCTATGAATTGAGGTTTATAGCTGAGAGAGCTCAGTTTATTTCTAAA | 22452 |
| QY | 90 ----- | 90 |
| Db | 22453 TTGATCGAGATCTCTTGCGGAGAAAGTACCTTTAGGGTTCTTTACTTAATTGAGAA | 22512 |

[illegible]

| | |
|--|--|
| REFERENCE | 1 (bases 1 to 1494) |
| AUTHORS | McCormick,S. |
| TITLE | Tomato mature anther specific protein (LAT61). |
| JOURNAL | Unpublished |
| REFERENCE | 2 (bases 1 to 1494) |
| AUTHORS | McCormick,S. |
| JOURNAL | Direct Submission |
| FEATURES | Submitted (27-JUN-2001) Plant Gene Expression Center, USDA/ARS/JC-Barkerley, 800 Buchanan St., Albany, CA 94710, USA Location/Qualifiers |
| source | 1. 1494 /organism="Lycopersicon esculentum" /mol_type="mRNA" /db_xref="taxon:4081" 21. 1022 /note="similar to protein encoded by Genbank Accession number AC025808" |
| CDS | /codon_start=1 /product="mature anther-specific protein LAT61" /protein_id="AAK71662.1" /db_xref="GI:14626761" /translation="MMWAGSPASSSSNAGAGGSGAGVGHPESGGGGGRPKPE RERNRRRRERRRAVAKIYTGAPACQNNYLPRICDNNNEUKALCTEGMIVPPQC TYRKCKRPMEIGTSTNTIPSSSRHISPSSEYASPIPSYOPSTSSFPSPSRF ANMLSHPSFLQNVPSSLPLRLISNSAIPVPSLPTRHPOTFNELTLAKSMFE NIPEHASAPASPTRVORFTPTTPTECHSESSTSDSCOMINFORKASNVNPPTPF VKPVPQPRPMIDTKDKSIDPEENTSVKRWEEGRIRHDVGFDELTLTSGNARI |
| BASE COUNT | 428 a 329 c 302 g 435 t |
| ALIGN | |
| Alignment Scores: | |
| Pred. No.: | 9.1e-41 Length: 1494 |
| Score: | 1060.00 Matches: 219 |
| Percent Similarity: | 72.54% Conservative: 32 |
| Best Local Similarity: | 63.29% Mismatches: 61 |
| Query Match: | 60.06% Indels: 34 |
| D: | 8 Gaps: 8 |
| US-09-995-938A-7 (1-336) x AF395901 (1-1494) | |
| Oy | 4 AepgylathnserThSerAlaalaalaaalaaalaaalaaalaa----- 21 |
| Dd | 39 GAATCACCAAGCATCTTCGCGCCGCCTCCGAGCTGTGTGAAGTGGAGTCCGAGTT 98 |
| Oy | 22 -----ArgarglyseProserTPARGlu 29 |
| Dd | 99 GCTTACCAGAAAGTGGTGGTGGTGGTGGTGGTGGAGAAGAAACCACTATGAGAGAA 158 |
| Oy | 30 ArggluanaanaaargargarglualrgrargargalavalalaaalyisleyThr 49 |
| Dd | 159 AGAGGAAATAACAGAGAAAGAGAGAGAGAGAGAGAGAGCTGTACTCTTAAGATTACT 218 |
| Oy | 50 Glyleuaaglainglyasptyrshuleupolyshicysaspashuanglualleu 69 |
| Dd | 219 GGTTAAAGCTCAAGCAAACTATPATCTTCGGAAGCACTGTGATPAACAATGAAGTCTT 278 |
| Oy | 70 LysalaleucysvalaiguilaaglyTtryvalaiguilaaprglyThrThryrArglys 89 |
| Dd | 279 AAAGCTCTTTTACTGAAAGCTGGTGGATCGTTGAACCTGATGTAACCACTTATCCCAA 338 |
| Oy | 90 GlyCyalyseProleuproglygluileaaaglyThrserserArgvalahpProtyrser 109 |
| Dd | 339 GGAATCAAGCCAAACCCGATGAGATTGAGAGCACTTCACAACAACATCAAGCCAATTCT 398 |
| Oy | 110 SergluhenglnserProleuSerSerxalapheginserProilleProserTyrginval 129 |
| Dd | 399 TCAGCGGATCCMAATCCCCCATCATCATCTTGTAGCCCAATTCATCATATCAAGCCA 458 |
| Oy | 130 SerProserSerSerSerPheProserProserArargglygluProashnaanaaSer 149 |
| Dd | 459 AGTCACACTTCCTCTTCTTCCCACATCTCATCTCGTCGTAAGGCC-----AACATGTTA 512 |
| Oy | 150 SerthrPhephePropheluarganaglyglyleProserSerleuProserleuaray 169 |


```

Db      513 TCACATCCATATCTTTCTCCAAAAT---GTCGTCCTTCATCCCTTCCTCATACGA 569
Oy      170 IleserAanSerCyProValThProProValSerSerProThSerIysanPolys 189
Db      570 AATCAAAACAGGCCCTTCTAATCCACCTTTCATCAACAACT---AGGCATCTTAAG 626
Oy      190 ProLeuProAanTrpGlnSerIleAlaLysGlnSerMetAlaIleAlaLysGlnSerMet 209
Db      627 CAACCTTCATTTAGAAACT-----TTGGCCAAAGAAATCAAG 665
Oy      210 AlaSerPheAanTyProPheTyAlaValSerAlaProAlaSerProThThsArgHis 229
Db      666 TTTCGTTAAACATCCCTTTCTTGTGCTTCAGCCCAAGCAACCACTAGGCTCAG 725
Oy      230 GlnPheHleThLeuAlaThrlleProGluCyAlaArgLysSerApsSerThrValaAp 249
Db      726 CGTTT---ACTCCTCAACTATACCCGAGTGATGAATCTGACTCATCTCACTTGAAT 782
Oy      250 SerGlyHleTrpIleSerPheGlnLysPheAlaGlnGlnGlnProPheSerAlaSerMet 269
Db      783 TCAGCCCACTGATCACTTCAAAAGTAT-----GCGTCAAT 821
Oy      270 ValProThSerProThPheAanLeuValLysProAlaProGlnGlnMetSerProAn 289
Db      822 GTTCACACCTTCCTCAACATTTATCTGTAAACCTGCTCCTCAGCCCTTCCTCAT 881
Oy      290 ThrAlaAlaPheGlnGlnIleGlyGlnSerSerGlnPheLysPheGlnAanSargVal 309
Db      882 GAT---ATGATCACAGACAGGTAAGACATAGACATTCGACTTGAATAATGATCAGTC 938
Oy      310 LysProTrpGluGlnGlnArgIleHleAanPValGlyMetGlnAanLysGlnLeuThrl 329
Db      939 AAGGATGGAAGGAGGTAAGGATCAAGATGAGATTCGATGATCTGAACTCACACTT 998
Oy      330 GlyAanGlyLysAlaArg 335
Db      999 GGAAGTGGCAATGCTCCG 1016

RESULT 11
AF372937 945 bp mRNA linear PLN 23-MAY-2001
LOCUS Arabidopsis thaliana Atlg19350/F18014_4 mRNA, complete cds.
DEFINITION AF372937.1 GI:13937166
ACCESSION AF372937.1 GI:13937166
VERSION FLI CDNA.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosa; eurosida II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 945)
Shum, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,
Bower, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayaishizaki, Y.,
Ishida, J., Jones, T., Kamuya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,
Nguyen, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,
Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 945)
Shum, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,
Bower, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayaishizaki, Y.,
Ishida, J., Jones, T., Kamuya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,
Nguyen, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,
Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Arabidopsis cDNA clones
Unpublished
TITLE JOURNAL
REFERENCE
AUTHORS

```

```

COMMENT
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN
Arabidopsis Full-Length cDNA) : Seki, M., Narusaka, M., Ishida, J.,
Seki, M., Kamuya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayaishizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Shum, P., Chen, H.,
Cheuk, R., Kim, C.J., Koesema, B., Meyers, M.C., Tracy, S.E., Banh, J.,
Bower, L., Chung, M.K., Goldsmith, A.D., Jones, T., Karlin-Neumann, G.,
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Southwick, A.,
Tang, C.C., Toriumi, M., Yamada, K., Yu, G., Davis, R.W., Theologis, A.,
and Ecker, J.R.

Shum, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
source
location/Qualifiers
1..945
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAF109-13-019(R12325)"
/note="ecotype: Columbia"
1..288
/note="ecotype: Columbia"
289..699
/note="unknown protein"
/codon_start=1
/product="Atlg19350/F18014_4"
/protein_id="AAK50077.1"
/db_xref="GI:13937167"
/translation="MSNAKQSMSTLNPYFAVSAPSPTHRQHPAPATPECEDSD
SSIVDSGHISFQKRAQQQPSASVSPFNVLVPPAPQSPPTAAIQEIQGSSEF
KFNBOVKPMGGRHIDVAMDELITLNGRAHS"
700..945

BASE COUNT 245 a 232 c 172 g 296 t
ORIGIN
3'UTR
BASE COUNT 245 a 232 c 172 g 296 t
ORIGIN

Alignment Scores:
Pred. No.: 7,79e-40 Length: 945
Score: 1035.00 Matches: 202
Percent Similarity: 90.04% Conservative: 6
Best Local Similarity: 87.45% Mismatches: 19
Query Match: 58.64% Indels: 4
Dn: 8 Gaps: 3

US-09-995-938a-7 (1-336) x AF372937 (1-945)
Oy 106 ThrProTySerSerGlnAanGlnSerProLeuSerSerAlaPheGlnSerProIlePro 125
Db 4 ACTCTTACTCTTCCCAACCAAGTCTCTTTTTCACACTTTGATAGCCCACTTA 63
Oy 126 SerTyGlnValSerSerSerSerSerPheProSerProSerArg--GlyGluPro 144
Db 64 TCTTCAACAGTACGCTCTCTCTTTCATTCACGAGCTCTTCGATGCTGATCA 123
Oy 145 AaAaAanMetSerSerThPhePheProPheLysAanGlyGlyLysProSer 164
Db 124 CACAAAT-----ATTCACACATCTTCCCTTCCAGGAAGTGATATTCCTTATCG 177
Oy 165 LeuProSerLeuArgIleSerAanSerCyAProValThProProValSerSerProTh 184
Db 178 CTTCCTCACTTGAATTTCAACAGCTGCTCTGTCACCTCCAGCTGATCCCACT 237
Oy 185 SerTyAanProTyProLeuProAanTyGlnSerIleAlaLysGlnSerMetAlaIle 204
Db 238 TCTAAGAACCCCAACCACTGCTCTTGGGATCTTTTACCAACCAATCATGTCATG 297
Oy 205 ---AlaLysGlnSerMetAlaSerPheAanTyProPheTyAlaValSerAlaProAla 223

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 20, 2003, 01:22:20 ; Search time 321 Seconds
(without alignments)
2825.580 Million cell updates/sec

Title: US-09-995-938a-7
Perfect score: 1765
Sequences: 1 MTSDGATSTGAAAAA.....IHVGMEDLELTGNGKARG 336

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.model/US0995938.rng -MINMATCH=0.1 -LOOPEL=0
-DB=N.GeneSeq.19Jun03 -OPMT=fbsbcp -SUFFIX=p2n.rng -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0995938.QCGN.1.1.312.rng -RUNAT 19122003.142606.4227 -NCPU=6 -ICPU=3
-NO MAP -LARGEJOINT -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N.GeneSeq.19Jun03.*

1: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2001.DAT.*
23: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2002.DAT.*
24: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2003.DAT.*
25: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2003.DAT.*

Prtd. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|--------|---------------|--------|-------------|----------------------|
| 1 | 1758 | 99.6 | 1501 | 21 AAC37477 | Arabidopsis thaliana |
| 2 | 1511.5 | 85.6 | 1359 | 21 AAC37137 | Arabidopsis thaliana |
| 3 | 1511.5 | 85.6 | 1418 | 21 AAC36495 | Arabidopsis thaliana |
| 4 | 825 | 46.7 | 881 | 24 AB212580 | Arabidopsis thaliana |
| 5 | 541 | 30.7 | 1237 | 21 AAC51043 | Arabidopsis thaliana |
| 6 | 541 | 30.7 | 1237 | 21 AAC42550 | Arabidopsis thaliana |
| 7 | 496.5 | 28.1 | 642 | 21 AAC42317 | Arabidopsis thaliana |
| 8 | 377 | 21.4 | 495 | 22 AAH87878 | Pepermint plant o |
| 9 | 371 | 21.0 | 433 | 25 ABX22943 | Human GDP-mannose |
| 10 | 321.5 | 18.2 | 284 | 25 ABX32572 | Mutational hot spo |
| 11 | 171.5 | 9.7 | 1050 | 22 AAD21685 | Human retinitis p |
| 12 | 171.5 | 9.7 | 2849 | 22 AAD21684 | Human retinitis p |
| 13 | 159.5 | 9.0 | 2800 | 22 AAD07131 | Human cDNA differe |
| 14 | 156.5 | 8.9 | 2805 | 22 AAD07130 | Human retinitis p |
| 15 | 153.5 | 8.7 | 227968 | 24 ABK83497 | Parallel detection |
| 16 | 149.5 | 8.5 | 2803 | 22 AAD07132 | DNA encoding novel |
| 17 | 148.5 | 8.4 | 2586 | 22 AAH21064 | Corn ear-derived p |
| 18 | 147.5 | 8.4 | 2582 | 23 AAS75052 | Drosophila melanog |
| 19 | 147.5 | 8.3 | 156 | 25 ABX88121 | S. avermilitis ave |
| 20 | 146.5 | 8.3 | 6521 | 23 ABH11153 | Streptomyces averm |
| 21 | 143.5 | 8.1 | 31422 | 21 AAH92302 | HSV L/ST region. |
| 22 | 143.5 | 8.1 | 31422 | 22 AAH92302 | Human immune/hema |
| 23 | 141.5 | 8.0 | 12001 | 16 AA076213 | Plasmodium falcipa |
| 24 | 141 | 8.0 | 6286 | 22 AAH75248 | P. falciparum tran |
| 25 | 140 | 7.9 | 9408 | 21 AAH70157 | M. capsulatus gene |
| 26 | 140 | 7.9 | 9606 | 15 AAQ67190 | Leishmania major p |
| 27 | 138.5 | 7.8 | 1428 | 24 ABQ91674 | DNA induced in hum |
| 28 | 138 | 7.8 | 1313 | 22 AAH74867 | Human EGR-1 cDNA. |
| 29 | 137.5 | 7.8 | 3109 | 13 AAQ24982 | Human EGR-1 coding |
| 30 | 137.5 | 7.8 | 3152 | 20 AAH64286 | Human EGR-1 DNA. |
| 31 | 137.5 | 7.8 | 3152 | 21 AAH64385 | Nucleotide sequenc |
| 32 | 137.5 | 7.8 | 3152 | 21 AAH64123 | Human cDNA encodin |
| 33 | 137.5 | 7.8 | 3152 | 21 AAH64123 | Human cDNA encodin |
| 34 | 137.5 | 7.8 | 3152 | 22 AAH64123 | Human lung tumor |
| 35 | 137.5 | 7.8 | 3152 | 22 AAH64123 | Nucleotide sequenc |
| 36 | 137.5 | 7.8 | 3152 | 22 AAH64123 | Gene #3724 used to |
| 37 | 137.5 | 7.8 | 3152 | 24 ABK84562 | Human benign prost |
| 38 | 137.5 | 7.8 | 3152 | 24 ABK84562 | cDNA encoding clon |
| 39 | 137.5 | 7.8 | 3152 | 24 ABK84562 | Colon adenocarcino |
| 40 | 137.5 | 7.8 | 3152 | 24 ABK84562 | Kidney cancer rela |
| 41 | 137.5 | 7.8 | 3152 | 24 ABK84562 | Human lung cancer |
| 42 | 137.5 | 7.8 | 3152 | 25 ACA10642 | Lung cancer therap |
| 43 | 137.5 | 7.8 | 3152 | 25 ABX99593 | DNA sequence enco |
| 44 | 137.5 | 7.8 | 3152 | 25 ABX10949 | |
| 45 | 137.5 | 7.8 | 3152 | 25 ABX10949 | |

ALIGNMENTS

RESULT 1
AAC37477
ID AAC37477 standard; DNM, 1501 BP.
XX AAC37477;
AC 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17521.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0128845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149378.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152263.
PR 07-SEP-1999; 99US-0152370.
PR 10-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0153770.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132483.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135529.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136582.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137228.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145513.
PR 27-JUL-1999; 99US-0145518.
PR 27-JUL-1999; 99US-0145519.
PR 28-JUL-1999; 99US-0145551.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149425.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139452.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139889.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0146388.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149729.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154039.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 23-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
 Pred. No.: 2,24e-25
 Score: 541.00
 Percent Similarity: 48.31%
 Best Local Similarity: 40.17%
 Query Match: 30.65%
 DB: 21
 Length: 1213
 Matches: 143
 Conservative: 29
 Mismatch: 64
 Indels: 120
 Gaps: 17

US-09-995-938a-7 (1-336) x AAC51043 (1-1213)

QY 23 ArglySPROSerTTPArgGluuArgLuanAmaAgaArgGluuArgArgala 42
 Db 163 AGAACGCGGACTTGAAGAGAGAGAGAACAAACGCGAGAGCGGAGAGACGAGG 222
 QY 43 ValAlaAlaValleIleTyrThrGlyLeuAgaAlaGlnGlyAspTyrAsnLeuProlyHis 62
 Db 223 ATTCGGGCTAGATCTTCGAGAGACTTAAGGATTCATGGAACCTTCAGGCTCTAAACAC 282
 QY 63 CysAspAsnAsnGluValLeuLysAlaLeuCysValAlaAlaGlyTyrValValGluGlu 82
 Db 283 TGGCAACAACATGAACTCTCAAGCTTTATGCAATGAAGCGCTTGAGCTGAAGAC 342
 QY 83 AspGlyThrThrTyrArgGlyGlyCysLysProLeuProGlyGluIleAlaGlyThrSer 102
 Db 343 GACGGAGACTACTACCGCAAGGAGATGCAACCAATG----- 378
 QY 103 SerArgValThrProTyrSer---SerGlnAsnGlnSerProLeuSerSerAlaPheGln 121
 Db 379 GATCGAATGGAAGCTCATGATGATGTTCTACTTACGCTAATGCTCATGCTCATGCAACAT 438
 QY 122 SerProIleProSerTyrGlnValSerProSerSerSerSerPheProSerProSerArg 141
 Db 439 AGCCCTCGTCTTCTTCAATCAATCCAGCCCTTGCTTTCATCATCTCCGAGTCTTCAACAC 498
 QY 142 ----GlyGluProAsnAsnAsnMetSerSerThrPhePheProPheLeuArgAsn--- 158
 Db 499 CCATTGGATGATGCTAAC-----TCACTAATCCATGCTGCTCAAGAACCC 543
 QY 159 ---GlyGlyIleProSerSerIleuProSerLeuArg---IleSerAsnSerCysProVal 176
 Db 544 TCTTCAAACTCACTCCAAAGCTTCCCTTTCATGGAATTCATATGAAGCCCTCCGCG 603
 QY 177 ThrProValSerSerProThrSerIleLysAsnProLysProLeuProAsnTrpIleuSer 196
 Db 604 ACTCCGCAATGGCT----- 618
 QY 197 IleAlaLysGlnSerMetAlaIleAlaLysGlnSerMetAlaSerPheAsnTyrProPhe 216
 Db 618 ----- 618
 QY 217 TyrAlaValSerAlaProIleSerProThrIleArgHisGlnPheIleThrLeuAlaThr 236
 Db 619 -----CGAAGCCCTACT---CGTGAACAA-----GTAACC 645
 QY 237 IleProGluCysAspGluSerAspSerSerThrValAspSerGlyHisTrpIleSerPhe 256
 Db 646 ATCCCT-----GACTCTGA---TGGCTCTCA--- 669
 QY 257 GlnLysPheAlaGlnGlnInProPheSerAlaSerMetValProThrSerProThrPhe 276
 Db 670 -----GGAATGCAAACTCCGAGAGCGGA-----CCGCTTCTCTACTTTC 711
 QY 277 AsnLeuVal----- 279
 Db 712 AGTTAGTTTCAAGAAACCGTTTTCGAAAGAGGCTTTAAATGGAGGATTTGTAAT 771
 QY 280 -----LysProAla---ProGlnGlnMet 286
 Db 772 TCACCAATGTGAGCTCCGAGCAAAAGTGAAACCTGCTCCAGCTAATTCCTGCTGTT 831
 QY 287 SerProAsnThrAlaAlaPheGlnGlnIleGlyGlnSerSerGluPheLysPhe----- 304
 Db 832 GATCGAAGCTCGATGTCGATGCTGATGCAAGCGCGAGCTTCCGTTTGGTTGTTGT 891
 QY 305 -----GluAsnSerGlnValLysProTyrGlyGlyGluArgGlnHis---Asp 319
 Db 892 AACGCAATGCTCGCAATGCAATGCGAAGCTTGGGAAAGGAGAAAGATACATGAGCA 951
 QY 320 ValGlyMetGluAspLeuGluLeuThrLeuGlyAsnGlyLysAlaArg 335
 Db 335 ----- 335

Db 952 TGTGTTGAGATGATTGAACTTACACTTGGAACACTGAAGACGAGA 999
 RESULT 6
 AAC42550
 ID AAC42550 standard; DNA; 1217 BP.
 XX
 AC AAC42550;
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35978.
 XX
 KW Hybridization assay; genetic mapping; gene expression control;
 KM protein identification; signal transduction pathway;
 XX metabolic pathway; promoter; termination sequence; ss.
 XX Arabidopsis thaliana.
 OS
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130074.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139456.

QY 60 ProLYAHlaCYaSPaSPaNSaNGluValleuLYaAlaLeuCYaValGluIaGluTYrPVal 79
 DB 288 CTTAAGCACTGGGATACACAGAGGTTCTTAAGCTCTGTCTCAAGCTGGTGGATC 348
 QY 80 ValGluIaSPaSPaThrThrTYrArgLYaGlyCYaLYaProLeuProGluIaLea 99
 DB 349 GTCGAAGACGATGGCACCACTTATCCAGGGGTTTAC-CCACACACATCAGATATTTCa 407
 QY 100 GlyThrSerSerArgValThrProTYrSerSerGlnaNSaGlnSerProLeuSerSera 119
 DB 408 GGAACCTCCACAACTTCAGACAAATTCATCATCAACCAAGTCACATCATCAGCT 467
 QY 120 PheGlnSerProIleProSerTYrGlnValaSerProSerSerSerPheProSerPro 139
 DB 468 TTTCACAGTCTCCACCTTCGACCGAGAGTCGCTCATCATCTTCCTTCCACACATC 527
 QY 140 SerArgGlyIuProaNSaNSaNSaMetSerSerThrPheProPheLeuArgaNS- 158
 DB 528 TCTCGCTATGAC--CGAAACCTTCTTCTATACCTTCTTCCGTTCTTCCACACATC 584
 QY 159 ---GlyGlyIleProSerSerleuProSerleuArgIleSerAnsSerCysProVal 176
 DB 585 GCTTCTTCGATTCGCTACCTTCCACCTTGAATATCCACAGTCGCTGTG 641
 RESULT 8
 AAH87878 standard; cDNA, 459 BP.
 AC AAH87878;
 DT 25-SEP-2001 (first entry)
 DE Peppermint plant oil gland expressed cDNA 234.
 KW Peppermint, plant oil gland cell; terpenoid essential oil; resin;
 KM genetic mapping; antisense suppression; recombinant expression; ss.
 OS Mentha x piperita.
 XX MO200153319-A1.
 XX 26-JUL-2001.
 PD 19-JAN-2001; 2001WO-US02567.
 PF 20-JAN-2000; 2000US-017264.
 PR (CROT/) CROTEAU R. B.
 PA (LANG/) LANGE B M.
 PA (WILD/) WILDUNG M R.
 PI Croteau RB, Lange BM, Wildung MR;
 XX WPI, 2001-488706/53.
 DR New nucleic acid molecules corresponding to mRNA molecules expressed in
 PT peppermint oil glands for enhancing expression of plant oil gland cell
 PT proteins -
 XX Claim 1; Page 165; 251pp; English.
 CC The invention relates to nucleic acid molecules (AAH87645-AAH88116) that
 CC correspond to all or part of a mRNA molecule expressed in plant oil
 CC gland cells, especially peppermint and plant oil glands that produce
 CC terpenoid essential oils and resins. The nucleic acids are useful for
 CC genetically mapping a plant genome for genes expressed in plant oil
 CC gland cells and to suppress (for example by antisense suppression) or
 CC enhance their expression (for example by genetically transforming a
 CC plant cell with a replicable expression vector that expresses one or more
 CC proteins naturally expressed in plant oil gland cells). The nucleic acids
 CC are also useful for recombinant expression of plant oil gland proteins
 CC required for terpenoid essential oil and/or resin production in bacterial

CC and/or yeast cells.
 XX SQ Sequence 459 BP; 1061 A; 141 C; 119 G; 93 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,53e-15 Length: 459
 Score: 377.00 Matches: 73
 Percent Similarity: 78.73% Conservative: 14
 Best Local Similarity: 61.86% Matches: 23
 Query Match: 21.36% Indels: 8
 DB: 22 Gaps: 1
 US-09-995-938A-7 (1-336) x AAH87878 (1-459)
 QY 23 ArgLYaSPaSPaThrThrTYrArgLYaGluValaNSaNSaArgArgLYaArgArgArg 42
 DB 130 AGCTGCTCCACATGAGAGAGCGGGAACAACAGCGAGGAGCGCGCCCGCGCG 189
 QY 43 ValAlaAlaLYaIleTYrThrGlyLeuArgAlaGlnGlyAspTYrAnsLeuProLYaHle 62
 DB 190 ATCGCGCGAAGATCTTCTCCGCTGAGAGATGTACGCACTCAAGCTCCCAAGCAC 249
 QY 63 CysAspAsNSaNSaGluValleuLYaAlaLeuCYaValGluIaGlyTYrPValValGlu 82
 DB 250 TCGACACACAAAGAGTGTCTCAAGCTCTTCCAAAGAGCTGGCTGGATCACAAGA 309
 QY 83 AspGlyThrThrTYrArgLYaGlyCYaLYaProLeuProGlyGluIaGlyThrSer 102
 DB 310 GACGCGACCACTTACAGAAAGGATGCAAGCTGTG----- 345
 QY 103 SerArgValThrProTYrSerSerGlnaNSaGlnSerProLeuSerSeraIaPheGlnSer 122
 DB 346 GAACGATGAGATATATAGGCTTCGGCACAGTCAGCTTGACCGTCATATCAACCAAGC 405
 QY 123 ProIleProSerTYrGlnValaSerProSerSerSerPheProSerProSer 140
 DB 406 CTTGAGGCTCTTTTAAACCAAGCTCTGATCTTCTTTGCTAGCCCTTC 459
 RESULT 9
 ABX22943
 ID ABX22943 standard; cDNA, 413 BP.
 XX ABX22943;
 AC ABX22943;
 DT 10-FEB-2003 (first entry)
 DE Human GDP-mannose 4,6-dehydratase (GM4, 6D) DNA #5000.
 XX Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
 KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
 KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
 KW complex carbohydrate; gene replacement therapy; immunosuppressive;
 KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
 KW antiasthmatic; vasotropic.
 XX Homo sapiens.
 OS
 XX US2002110548-A1.
 PN 15-AUG-2002.
 PD 11-JUN-2001; 2001US-0878574.
 PF 22-NOV-1996; 96US-0753233.
 PR 03-DEC-1997; 97US-0984246.
 PR 09-SEP-1998; 98US-0148674.
 PR 14-JUN-1999; 99US-0333177.
 XX (GENW) GENETICS INST INC.
 PA Sullivan F, Kriz R, Kumar R;
 PI WPI, 2003-066673/06.
 DR

[illegible]

| | | |
|----|---|---|
| | KW | cellular fucosylation; glycoconjugate fucosylation; transplant rejection |
| | KM | arthritis; asthma; sepsis; reperfusion injury; stroke; infection; |
| | KW | complex carbohydrate; gene replacement therapy; immunosuppressive; |
| | KM | antiinflammatory; antiarthritic; antibacterial; cerebroprotective; |
| | KW | antithrombotic; vasotrophic. |
| | XX | Homo sapiens. |
| | XX | US2002110548-A1. |
| | PN | |
| | PD | |
| | PP | 15-AUG-2002. |
| | XX | |
| | PF | 11-JUN-2001; 2001US-0878574. |
| | XX | |
| | PR | 22-NOV-1996; 96US-0753233. |
| | PR | 03-DEC-1997; 97US-0964246. |
| | PR | 09-SEP-1998; 98US-0149674. |
| | PR | 14-JUN-1999; 99US-0333177. |
| | XX | |
| | PA | (GEMV) GENETICS INST INC. |
| | PB | |
| | PI | Sullivan F, Kriz R, Kumar R; |
| | XX | |
| | XX | WPI; 2003-066673/06. |
| | XX | |
| | PT | New composition comprising GDP-mannose 4,6-dehydratase (GM4, 6D) |
| | PT | peptide, for manufacturing complex carbohydrates, or as targets for |
| | PT | screening GM4, 6D antagonists for treating e.g. arthritis, or transplant |
| | PT | rejection - |
| | PS | |
| | PS | Disclosure; SEQ ID NO 14631; 6pp; English. |
| | XX | |
| | CC | The invention relates to a composition comprising a human GDP-mannose |
| | CC | 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying |
| | CC | GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation |
| | CC | in a mammalian subject and for treating or ameliorating diseases affected |
| | CC | by the level of cellular fucosylation or diseases affected by the |
| | CC | fucosylation of glycoconjugates. These diseases include arthritis, |
| | CC | transplant rejection, asthma, sepsis, reperfusion injury, stroke or |
| | CC | infection. The GM4,6D peptide or a polynucleotide encoding it is also |
| | CC | useful for manufacturing complex carbohydrates and as targets for |
| | CC | screening small molecule antagonists of the activity of the enzyme. The |
| | CC | polynucleotide is useful in developing an assay for defects in the |
| | CC | enzyme, as well as in gene replacement therapy. Sequences |
| | CC | ABX17942-ABX17944 and ABX17947-ABX3716 represent DNA molecules encoding |
| | CC | human GM4,6D peptides of the invention. |
| | CC | Note: The sequence data for this patent did not form part of the printed |
| | CC | specification but was obtained in electronic format directly from USPTO |
| | CC | at seqdata.uspto.gov/sequence.html. |
| | XX | |
| | XX | Sequence 284 BP; 95 A; 62 C; 89 G; 38 T; 0 other; |
| | XX | |
| | SO | |
| | Alignment Scores: | |
| | Pred. No.: | 2,73e-12 Length: 284 |
| | Score: | 321.50 Matches: 66 |
| | Percent Similarity: | 81.91% Conservative: 11 |
| | Best Local Similarity: | 70.21% Mismatch: 5 |
| | Query Match: | 18.22% Indels: 12 |
| | DB: | Gaps: 1 |
| | US-09-995-938A-7 (1-336) x ABX35572 (1-284) | |
| OY | 4 | ASPGYALATHYSErThSeRlaIAlaIAlaIAlaIAlaIAlaIAlaIAlaIArG-Ar 23 |
| Db | 17 | GACGGAGCAACTCG-----GCCGCAGCAGCCGGAAg 49 |
| OY | 23 | GLYFPROSeRTTPAdGGu-ArgGUaaMaMaAtyArGrG-GluHrgARGaRGaRGAla 42 |
| Db | 50 | AAAGCCGTCGTGAAGAAGGCAACACAGAAAGAGAAAAGACGMAAGAGCA 109 |
| OY | 43 | VAlAlaAlaLySIlEtYrTTrGlYLeuArGAlaInGlyAsPTrzAnleuPePolYShIs 62 |
| Db | 110 | ATAGACGCAAAATATCTACTCAAGACTTCAGACTCAGGGAACTTCACATTGCCAAAGCAT 169 |

| ID | Accession | Gene | Feature | Location/Qualifiers |
|-------------|---|------|---------|---|
| 63 | CYAAAPAAASBGLVALLLEULYSALALEUCYVALGLIATTPVALVAILGUGLU | 82 | | |
| 170 | TGCGACACACACAGCGTTCTGAAACCTCTTGCGCAAGACTGGTTGGCGTAGAAGAA | 225 | | |
| 83 | AAPGlyThrThrTyrArgLysGlyCysAlaProLeu | 94 | | |
| 230 | GACGGACCACTTATCGCAAGGGTTGCAACGGCTC | 265 | | |
| RESULT 11 | | | | |
| AAD21685/c | AAD21685 standard; DNA; 1050 BP. | | | |
| AAD21685; | | | | |
| 28-JAN-2002 | (first entry) | | | |
| DE | Mutational hot spot sequence of human RPGR exon ORF15 DNA. | | | |
| XX | | | | |
| XX | Human; mutation; retinitis pigmentosa GTPase regulator; RPGR; | | | |
| XX | genotyping; open reading frame; ORF; X-linked retinitis pigmentosa; | | | |
| KW | XLR; gene therapy; screening; forensic analysis; ds. | | | |
| XX | | | | |
| OS | Homo sapiens. | | | |
| XX | | | | |
| Key | Location/Qualifiers | | | |
| FT | FT | FT | FT | 1..73 |
| FT | FT | FT | FT | /tag= a |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 93..94 |
| FT | FT | FT | FT | /tag= b |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 102 |
| FT | FT | FT | FT | /tag= c |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 257..258 |
| FT | FT | FT | FT | /tag= d |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 262..263 |
| FT | FT | FT | FT | /tag= e |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 283..284 |
| FT | FT | FT | FT | /tag= f |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 299..302 |
| FT | FT | FT | FT | /tag= g |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 379 |
| FT | FT | FT | FT | /tag= h |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 414 |
| FT | FT | FT | FT | /tag= i |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 482 |
| FT | FT | FT | FT | /tag= j |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 507 |
| FT | FT | FT | FT | /tag= k |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 524..526 |
| FT | FT | FT | FT | /tag= l |
| FT | FT | FT | FT | /note= "Benign or polymorphic sequence variant" |
| FT | FT | FT | FT | 587..588 |
| FT | FT | FT | FT | /tag= m |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 620..621 |
| FT | FT | FT | FT | /tag= n |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 751 |
| FT | FT | FT | FT | /tag= o |
| FT | FT | FT | FT | /note= "Mutational hot spot" |
| FT | FT | FT | FT | 849..853 |
| FT | FT | FT | FT | /tag= p |

```

FT FT misc_feature /note= "Mutational hot spot"
FT FT 907..908 /+tag=
FT FT /tag= G
FT FT /note= "Mutational hot spot"
FT FT misc_feature /tag= 954
FT FT /note= "Mutational hot spot"
PN MO200177380-A2.
XX
XX 18-OCT-2001.
PD
XX 10-APR-2001; 2001WO-GB01622.
XX
XX 10-APR-2000; 2000GB-0008801.
PA (MEDI-) MEDICAL RES COUNCIL.
XX
XX Wright A;
XX
XX WPI; 2001-663057/76.
XX
XX Diagnosing disease or predisposition to disease, associated with
XX disease causing mutations in retinitis pigmentosa GTPase regulator gene
XX by genotyping ORF15 of the gene, and determining presence of mutations
XX
XX
XX Claim 6, Page 90, 100pp, English.
XX
XX The present invention relates to a method for diagnosing disease or
XX predisposition to a disease, associated with a disease causing
XX mutations in a retinitis pigmentosa GTPase regulator (RPGR) gene
XX involves genotyping a RPGR gene, and determining whether the genotype
XX comprises a disease causing mutations, where the risk genotype is
XX present within open reading frame (ORF)15 of the RPGR gene. The method
XX is useful for detecting a certain disease state e.g., X-linked
XX retinitis pigmentosa (XLRP). The kit is useful for detecting and
XX measuring disease causing mutations in biological fluids and tissues
XX and for localising mutation in tissues. The mutant RPGR gene is useful
XX in gene therapy techniques and for screening agents capable of
XX affecting the expression of the sequences and/or the biological
XX activity of mutant RPGR. They are preferably useful for identifying
XX agonists and antagonists of RPGR. The mutant RPGR gene is also useful
XX in identification of potential pharmaceutical targets in high
XX throughput screening assays and forensic analysis. The present sequence
XX is the mutational hot spot sequence of human RPGR exon ORF15 DNA.
XX
XX Sequence 1050 BP, 428 A, 15 C, 595 G, 12 T, 0 other:
XX
XX
XX Alignment Scores:
XX Pred. No.: 0.0345 Length: 1050
XX Score: 171.50 Matches: 69
XX Percent Similarity: 44.09% Conservative: 28
XX Best Local Similarity: 31.36% Mismatches: 77
XX Query Match: 9.72% Indels: 47
XX DB: 22 Gaps: 10
XX
XX US-09-995-938A-7 (1-336) x AAD21685 (1-1050)
XX
XX 93 ProLeuProGlyGluIleAlaGlyThySerSerArgValThrProTyr---SerSerGln 111
XX |||||
XX 735 CCTCTTCCCTCCCTCCATCTCTCT-TCNCCTCTTCTCCCTCCCTCTTCATCTCCCT 6777
XX |||||
XX 112 AsnGlnSerProLeuSerSerAlaIleAsnGlnSerProIleProSerTyrGlnValSerPro 111
XX |||||
XX 676 TCCCTCTTCTCT---TCTCTCTCTCCCTCCCTCTCCCTCTCTCTCTCC-----TCTCTCT 629
XX |||||
XX 132 SerSerSerSerPheProSerProSerArgGlyGluProAsnAsnAsnMetSerSerThr 151
XX |||||
XX 628 TCTCTCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 569
XX |||||
XX 152 PhePheProPheLeuArgAsnGlyGlyIleProGlnSerLeuProSerSerLeuArgIleSer 171
XX |||||

```

[illegible]

| | | |
|----|---|---|
| PT | misc_feature | /note= "Mutational hot spot" |
| FT | | 694..708 |
| FT | | /*tag= j |
| FT | misc_feature | /note= "Benign or polymorphic sequence variant" |
| FT | | 769 |
| FT | | /*tag= k |
| FT | misc_feature | /note= "Mutational hot spot" |
| FT | | 804 |
| FT | | /*tag= l |
| FT | misc_feature | /note= "Mutational hot spot" |
| FT | | 872 |
| FT | | /*tag= m |
| FT | misc_feature | /note= "Mutational hot spot" |
| FT | | 897 |
| FT | | /*tag= n |
| FT | misc_feature | /note= "Mutational hot spot" |
| FT | | 977..978 |
| FT | | /*tag= o |
| FT | misc_feature | /note= "Mutational hot spot" |
| FT | | 1010..1011 |
| FT | | /*tag= p |
| FT | misc_feature | /note= "Mutational hot spot" |
| FT | | 1067..1087 |
| FT | | /*tag= q |
| FT | misc_feature | /note= "Benign or polymorphic sequence variant" |
| FT | | 1141 |
| FT | | /*tag= r |
| FT | misc_feature | /note= "Mutational hot spot" |
| FT | | 1166..1186 |
| FT | | /*tag= s |
| FT | misc_feature | /note= "Benign or polymorphic sequence variant" |
| FT | | 1239..1243 |
| FT | | /*tag= t |
| FT | misc_feature | /note= "Mutational hot spot" |
| FT | | 1297..1298 |
| FT | | /*tag= u |
| FT | misc_feature | /note= "Mutational hot spot" |
| FT | | 1307..1318 |
| FT | | /*tag= v |
| FT | misc_feature | /note= "Benign or polymorphic sequence variant" |
| FT | | 1321..1332 |
| FT | | /*tag= w |
| FT | misc_feature | /note= "Benign or polymorphic sequence variant" |
| FT | | 1344 |
| FT | | /*tag= x |
| FT | misc_feature | /note= "Mutational hot spot" |
| FT | | 2818..2823 |
| FT | | /*tag= y |
| XX | | |
| XX | polyA_signal | |
| XX | | |
| XX | WO200177380-A2. | |
| XX | | |
| XX | 18-OCT-2001. | |
| XX | | |
| XX | 10-APR-2001; 2001WO-GB01622. | |
| XX | | |
| XX | 10-APR-2000; 2000GB-0008801. | |
| XX | | |
| XX | (MEDI-) MEDICAL RES COUNCIL. | |
| XX | | |
| XX | Wright A; | |
| XX | | |
| XX | WPI; 2001-663057/76. | |
| XX | | |
| XX | P-PSDB; AAE13147. | |
| PT | | |
| PT | Diagnosing disease or predisposition to disease, associated with | |
| PT | disease causing mutations in retinitis pigmentosa GTPase regulator gene | |
| PT | by genotyping ORF15 of the gene, and determining presence of mutations | |
| PT | - | |
| XX | | |
| XX | Disclosure; Fig 2; 100pp; English. | |
| PS | | |
| XX | | |
| CC | The present invention relates to a method for diagnosing disease or | |
| CC | predisposition to a disease, associated with a disease causing | |


```
Qy 137 ProSerProSerArgGlyGluProAsnAsnMetSerSer----- 150
Db 81789 CCTTCCCCAGAGCCGCGACCTCGGGAGAGCCCGGAGCCGCGCTTAGGGG 81848
Qy 151 -----ThrPhePheProPheLeuArgAnGlyGlyIleProSer 163
Db 81849 GCGACCGCTCCGCCCCCTCGGACCTGGGGCCCCCTTCCCTGAG-----GCGACGGCTTC 81902
Qy 164 SerLeuProSerLeuArgIleSerAsnSerCysProValThrProProValSerSerPro 183
Db 81903 TCCCGG-----GCGGGCGGCTGCTCCACCCCGCGCTCTTACGGTC 81947
Qy 184 ThrSerIlyAsnProIlyProIlyProIlyProIlySerIleAlaIlyGlnSerMetAla 203
Db 81948 GCC-----CCGGGGCCCTGGCCA-----AGCCCTGGCCCTGGCGCGG 81989
Qy 204 IleAlaIlyGlnSerMetAlaSerPheAsnTyProPheTyAlaValSerAlaProAla 223
Db 81990 GCGCCCGACCTAGCGTCTTAGGGTGGCGGGCCCGGCGCTTCCGAGTTCTCCAGC 82049
Qy 224 SerPro-----ThrIleArgHleGln-----Phe 231
Db 82050 GCGCTTTGGGTTCCCTCGGAGAGAGCTGGGCTTGGGTTCCCAAGTCCAGAGCTTC 82109
Qy 232 His-----ThrLeuAlaThrIleProGluCysAspGluSerAspSer 245
Db 82110 CACGACTTTGGCCAACATGAAAGCTTGCCTCCACCCAGGACCTCGCTGGTCTCC 82169
Qy 246 SerThrValAspSer-----GlyHleTyrIleSerPheG1 257
Db 82170 TCACCTCTGTTCTTCTTGGCTCCGGGATCTTAGTGTGGTCAATGGGGTCTCTTCTT 82229
Qy 257 nIyAsPheAlaGlnGlnGln-----ProPheSerAlaSerMetValProThrSerPr 274
Db 82230 CCCACTTCTCGGCGCGGAAATACCTTCTTACAGAGACTGCTTGTCTTCCCCGTC 82289
Qy 274 oThrPheAsnLeuValIlyProAlaProGlnGlnMetSerProAsnThrAlaIlyPheG1 294
Db 82290 C-----AGGCCAGGCCCTTCTCAGCTCATGCCCTCCGTCAGCA---GC 82331
Qy 294 nGluIleGlyGlnSerSerGluPheIlyPheGluAsnSerGlnValIlyS 310
Db 82332 TCCTGGCCACAGCTCTCAGACTTGTCTTAAAAAACAAACAA 82380
```

Search completed: December 20, 2003, 05:50:38
Job time : 372 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 20, 2003, 01:10:37; Search time 2483 Seconds

(without alignments)
3288.886 Million cell updates/sec

Title: US-09-995-938A-7
Perfect score: 1765
Sequence: 1 MTSQATSTSAIAAAAAAAAA.....IHDVMEDELITLNGKARG 336

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-Q/cgn2.1/USPTO.spool/US0995938/runat.19122003.142606.4249/app.query.fasta_1.519
-DB=EST -OPMT=fastcap -SUPFLX=p2n.ref -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0995938 @cgn 1 1 2810 @runat.19122003.142606.4249 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBSTR -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rdg:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 913.5 | 51.8 | 772 | 13 CB292898 | CB292898 UCRC001.0 |
| 2 | 901 | 51.0 | 796 | 14 BU238047 | BU238047 DB01.14F0 |
| 3 | 885.5 | 50.2 | 789 | 10 BG127474 | BG127474 EST473120 |
| 4 | 872.5 | 49.4 | 765 | 10 BG126322 | BG126322 EST471968 |
| 5 | 871 | 49.3 | 768 | 12 BG887699 | BG887699 EST513550 |
| 6 | 857 | 48.6 | 585 | 9 AV551486 | AV551486 AV551486 |
| 7 | 838.5 | 47.5 | 762 | 12 B1175815 | B1175815 EST516910 |
| 8 | 826.5 | 46.8 | 766 | 10 BG593827 | BG593827 EST492505 |
| 9 | 778 | 44.1 | 658 | 10 BG130895 | BG130895 EST463787 |
| 10 | 777.5 | 44.1 | 643 | 14 CA934625 | CA934625 MTUSTS.P2 |
| 11 | 776.5 | 44.0 | 619 | 9 AV440980 | AV440980 AV440980 |
| 12 | 763.5 | 43.3 | 941 | 28 B12579 | B12579 F22H5-T7 IG |
| 13 | 759.5 | 43.0 | 648 | 9 AV545003 | AV545003 AV545003 |
| 14 | 752 | 42.6 | 602 | 29 B2476824 | B2476824 BOMBH01TR |
| 15 | 746 | 42.3 | 639 | 9 AV822043 | AV822043 AV822043 |
| 16 | 743 | 42.1 | 1001 | 12 BG840094 | BG840094 GC01_0200 |
| 17 | 732.5 | 41.5 | 864 | 14 CB893187 | CB893187 EST645979 |
| 18 | 732 | 41.5 | 744 | 12 B1921662 | B1921662 EST541565 |
| 19 | 730.5 | 41.4 | 773 | 28 BH673471 | BH673471 BOHXG46TR |
| 20 | 729 | 41.3 | 685 | 9 AV782667 | AV782667 AV782667 |
| 21 | 720.5 | 40.8 | 840 | 14 CB892700 | CB892700 EST645492 |
| 22 | 719.5 | 40.8 | 729 | 13 BU027742 | BU027742 QHG7H23.Y |
| 23 | 716.5 | 40.6 | 717 | 13 BU027177 | BU027177 QHG2E02.Y |
| 24 | 711.5 | 40.3 | 708 | 10 BF480563 | BF480563 LO-251473 |
| 25 | 709.5 | 40.2 | 724 | 13 BU024180 | BU024180 OHF16J24. |
| 26 | 696.5 | 39.5 | 717 | 13 BU025606 | BU025606 OHG10E24. |
| 27 | 692 | 39.2 | 553 | 10 BE450956 | BE450956 EST401843 |
| 28 | 689 | 39.0 | 658 | 13 BQ138490 | BQ138490 NF003G08P |
| 29 | 687 | 38.9 | 801 | 12 B1433025 | B1433025 EST535786 |
| 30 | 684 | 38.8 | 590 | 10 BE343539 | BE343539 EST408701 |
| 31 | 680.5 | 38.6 | 766 | 29 B2476333 | B2476333 BOMBH01TF |
| 32 | 673.5 | 38.2 | 610 | 10 BE803614 | BE803614 BR61B05.Y |
| 33 | 671 | 38.0 | 817 | 14 CB893419 | CB893419 EST646211 |
| 34 | 660 | 37.4 | 614 | 9 AW317351 | AW317351 B94BD10.Y |
| 35 | 655.5 | 37.1 | 677 | 9 AW299184 | AW299184 EST305994 |
| 36 | 654 | 37.1 | 625 | 10 BE022315 | BE022315 sm73c12.Y |
| 37 | 649 | 36.8 | 511 | 10 BF277350 | BF277350 GA_Eb002 |
| 38 | 645.5 | 36.6 | 629 | 13 BQ407744 | BQ407744 GA_Ed000 |
| 39 | 644 | 36.5 | 579 | 12 BM062278 | BM062278 KS01040G0 |
| 40 | 635.5 | 36.0 | 590 | 9 AV822219 | AV822219 AV822219 |
| 41 | 626 | 35.5 | 511 | 28 BH597004 | BH597004 BOCBP74TF |
| 42 | 626 | 35.5 | 539 | 9 AV442375 | AV442375 AV442375 |
| 43 | 623.5 | 35.3 | 679 | 9 AW689777 | AW689777 NF024C07S |
| 44 | 622 | 35.2 | 676 | 13 BQ917002 | BQ917002 QHB1H05.Y |
| 45 | 620 | 35.1 | 580 | 14 CD481007 | CD481007 eca01-6ms |

ALIGNMENTS

RESULT 1
CB292898
LOCUS
DEFINITION
albedo cDNA library Citrus sinensis
ACCESSION
CB292898
VERSION
CB292898.1
KEYWORDS
SOURCE
ORGANISM
Citrus sinensis
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

FEATURES
source

Location/Qualifiers
1. 796
/organism="Descurainia sophia"
/mol_type="mRNA"
/db_xref="taxon:89411"
/clone="De01_14F06"
/cissue_type="leaf, stem"
/dev_stage="1-month seedlings - 1 cm tall - 8 leaf"
/clone_lib="De01_AARC_ECORC_cold_stressed_Rixweed_seedlin
gs"
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: Eco RI;
Site 2: Xho I; Plants were grown for 1 month at 200C/16
hrs light/day (average 8 leaves, 1 cm tall, weight
0.02g/plant). Then they were exposed to 20C, 12 hrs
light/day, for 1 week. Library prepared by C. Fliche using
Stratagene kit."

BASE COUNT 221 a 150 c 211 g 207 t 7 others

Alignment Scores:

Pred. No.: 9.08e-58 Length: 796
Score: 901.00 Matches: 178
Percent Similarity: 87.56% Conservative: 5
Best Local Similarity: 85.17% Mismatches: 22
Query Match: 51.05% Indels: 4
DB: 13 Gaps: 3

US-09-995-938a-7 (1-336) x BU238047 (1-796)

QY 129 ValSerProSerSerSerSerPhePheProSerProSerArg---GlyGluProAlaAlaAsn 147
DB 796 GTCAGTCATCATGCTTCATCATCTCCAGTCCTTCGAGTGGCAGCCACACACAC--- 740
QY 148 MetSerSerThrPhePheProPheLeuArgAsnGlyGlyLeuProSerSerLeuProSer 167
DB 739 ---ATCTCCACAAATCTCCCGCTTCCTCAGAAATGGCGAAATCCTTCATGCTTCGCG 683
QY 168 LeuArgGlySerAsnSerCyPProValThrProProValSerSerProThrSerLeuAsn 187
DB 682 CTTAGGATCTCAACACATGCTCCAGTCATCCACAGCTTCATCTCCAAATTTCTAAAC 623
QY 188 ProLeuProLeuProAlaThrGlySerLeuAlaSerMetAlaAla---AlaLys 206
DB 622 CCAAAACCAATGCCCCACTGGAGCTTTTACCAACATCCATGCGCATGGCTGCTAA 563
QY 207 GlnSerMetAlaSerPheAsnThrProPheThrAlaValSerAlaProAlaSerProThr 226
DB 562 CAATCATGATCATCTTAACTATCCATTTACCCCGATTCGACCTGCCATGCCAC 503
QY 227 HisArgHisGlnPheHisThrLeuAlaSerThrLeuProGluCyAspGluSerAspSerSer 246
DB 502 CATCATGCGCAGTCCATGCGCCCGCTACTATTCCTGATGTGATGATGATGATCTTCC 443
QY 247 ThrValAspSerGlyHisThrPheSerPheGlnLysPheAlaGlnGlnGlnProPheSer 266
DB 442 ACTGTGACTCGGCTCATGTGATAGCTTTCAAAAGTTCTCAACAGCCATTCATGTC 383
QY 267 AlaserMetValProThrSerProThrPheAsnLeuValLysProAlaProGlnGlnMet 286
DB 382 GGCTCTCTCTGCCCCACTCTCCCAATCTCAATTAAGAAACCCGCMCTCCACAAATTA 323
QY 287 SerProAsnThrAlaAlaPheGlnGlnLysGlnSerSerGluPheLysPheGluAsn 306
DB 322 TCTCCAAACACTGCTGCAACCAAGAGATGTGTCAAACTCGGAATTTAAATTCGAGAAC 263
QY 307 SerGlnValLysProThrProGlnGlyGluArgLysHisAspValGlyMetGluAspLeuGlu 326
DB 262 AGCCAAATTAAGCATGGAGAGAGAGATCATGATGCGCTATGAGAGATCTAGAG 203
QY 327 LeuThrLeuGlyAsnGlyLysAlaArg 335
DB 202 CTCACACTTGGAAACGGCAAGCTCTGT 176

RESULT 3
BG127474
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BG127474 789 bp mRNA linear EST 31-JAN-2001
EST473120 tomato shoot/meristem Lycopersicon esculentum cDNA clone
CFOF16J21 5' sequence, mRNA sequence.
BG127474
BG127474.1 GI:12627662
EST.
Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 789)
van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,
Hansen, C., Romling, C. and Tanksley, S.
Generation of ESTs from tomato shoot/meristem tissue
Unpublished
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.
Location/Qualifiers
1. 789
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CFOF16J21"
/cissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/clone_lib="tomato shoot/meristem"
/note="Vector: Bluescript SK(-); Site 1: EcoRI; Site 2:
XhoI. Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."

BASE COUNT
ORIGIN

218 a 186 c 160 g 225 t

Alignment Scores:

Pred. No.: 1.3e-56 Length: 789
Score: 885.50 Matches: 181
Percent Similarity: 76.62% Conservative: 32
Best Local Similarity: 65.11% Mismatches: 46
Query Match: 50.17% Indels: 19
DB: 10 Gaps: 7

US-09-995-938a-7 (1-336) x BG127474 (1-789)

QY 33 AsnArgArgArgGluArgArgArgAlaValAlaValLysIleThrGlyLeuArg 52
DB 3 AATTAG 62
QY 53 AlaGlnGlyAspGlyAsnLeuProLysHisCysAspAsnAsnGlyValLeuValLeu 72
DB 63 GCACAGGGGAGATTATATCTTCTCAACATTTGATTAACAATGAGGTTTGAAGCTCTT 122
QY 73 CysValAlaLysGlyThrProValGluGluAspGlyThrThrThrArgLysGlyCysLys 92
DB 123 TGTGTTAAGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 182
QY 93 ProLeuProGlyGluIleAlaGlyThrSerSerArgValThrProThrSerSerGlnAsn 112
DB 183 CCAACTCCATGAGAGATTGGAGGACCTTCAGCCAAATTAAGCCAAAGTTCTTCAAGAAAT 242
QY 113 GlnSerProLeuSerSerAlaPheGlnSerProThrProThrProThrProThrProThr 132
DB 243 CCAAGTCTCTCCCTTCATCTTGTGATGAGCCGATTCATCTTCAAGTATGATCAACA 302
QY 133 SerSerSerPheProSerProSerArgGlyGluProAlaAsnAsnMetSerSerThrhe 152

Db 303 CCCTGCTGCTTCCCAAGTCACATCTGCTGATGTGCT-----AACATGTCGTCACATGCA 356
 Oy 153 PheProPheLeuAaGaaGlyGlyLeuProSerSerLeuProSerLeuAaGllSerAen 172
 Db 357 TTTGCAATTTCATC---CATAGTTCCATCTCCCTTGCTGCTACCAACCATTAGCAATATCAAC 413
 Oy 173 SerGysProValAlaThrProProValaSerSerProThrSerLysAanProLysProLeu 192
 Db 414 AGTGCCTCTGTACACCAACCTCTTTCATCACCACT--AGAGTCCCTTAAGCAGATATTT 470
 Oy 193 AenTpgJuserLleAlaLysGlnSerMetAlaLeuAlaLysGlnSerMetAlaSerPhe 212
 Db 471 AATCTTGAGACT-----TTGGCTAGAGACTTATGTCTGCTCTGA 509
 Oy 213 AenTyTrProPheTyAlaValaSerAlaProAlaSerProThrHisArgHisGlnPheHis 232
 Db 510 AATATCCCTTCTTGTGCTGCTTGCACCCCACTACCCCACTGAGTCAGACGATTC-- 566
 Oy 233 ThrLeuAlaThrLeuProGlnCysAerPglUserAaPserSerThrValaSerSerLys 252
 Db 567 ACTCTGCTGACAAATACCAAGAGTGTACGAATGTGATATCATCCACATGATATTTGGCAG 626
 Oy 253 TrpLleSerPheGlnLysPheAlaGlnGlnGlnGlnProPheSerAlaSerMetValProThr 272
 Db 627 TGGAGAGACTTTCAAAGATGCCA-----GCCAAATGGATCTCTACT 668
 Oy 273 SerProThrPheAanLeuValLysPheAlaProGlnGlnMetSerProAanThrAla 292
 Db 669 TCTCGCACTTTTAATCTTATTAAGCTGTAGCTCAGAGAAATCTTCTTAATGATATGATC 728
 Oy 293 PheGlnGlnLleGlyGlnSerSerGlnPheLysPheGlnAanSerGlnValLys 310
 Db 729 ATGCACAAAG--GGTAAGAGCAATTGAATTTGACTTGAAGAAATGATCAGATTAAAG 779

 RESULT 4
 BGI26322
 LOCUS BGI26322
 DEFINITION EST171868 tomato shoot/meristem Lycopersicon esculentum cDNA clone
 CTOP12A3 5' sequence, mRNA sequence.
 ACCESSION BGI26322
 VERSION BGI26322.1 GI:12626510
 KEYWORDS EST.
 SOURCE Lycopersicon esculentum (tomato)
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 765)
 van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Uterback,T.,
 Hansen,C., Roning,C. and Tankley,S.
 Generation of ESTs from tomato shoot/meristem tissue
 Unpublished
 TITLE JOURNAL
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>.
 location/Qualifiers
 1..765

ORIGIN

| | |
|------------------------|----------|
| Alignment Scores: | |
| Pred. No.: | 1,18e-55 |
| Score: | 873.50 |
| Percent Similarity: | 76.92% |
| Best Local Similarity: | 65.20% |
| Query Match: | 49.43% |
| DB: | 10 |
| | |
| Length: | 765 |
| Matches: | 178 |
| Conservative: | 32 |
| Mismatches: | 44 |
| Indels: | 19 |
| Gaps: | 7 |

US-09-995-938A-7 (1-336) x BG126322 (1-765)

| LOCUS | DEFINITION | LOCUS | DEFINITION |
|------------|---|------------|---|
| Db | 3 AsnArgArgArgGluuArgGArgGArgAValAlaIleAlaIleValIleTyrThrGluPheArg 52 | Db | 3 AAATGAGGAGGAGAAAGGAGAGAGGCAATAGGCACTTAAGATTTTAGTGGATTAAAGA 62 |
| Qy | 53 AlaGlnGlyAspTyrAsnLeuProLysHisCysAspAsnAsnGluValLeuLysAlaLeu 72 | Qy | 53 CysValAluAlaGlyTyrPValValGluGluAspGlyThrTyrArgGlyGlyCysLys 92 |
| Db | 63 GCACAGGGGAATTATATCTTCCTTAACAATTGGATGAATAGAGGTTTGACAGCTCTT 122 | Db | 123 TGTGTTGAACTGGATGATGATTGTTGAGCCGTGAATGAATCTATTAAAGAAAGGATGACGG 182 |
| Qy | 93 ProLeuProGlyGluIleAlaGlyThrSerSerArgValThrProLysSerSerGluAsn 112 | Qy | 93 GlnSerProLeuSerSerAlaPheGlnSerProIleProSerTyrGlnValSerProSer 132 |
| Db | 183 CCAACTCCAAATGGAGATTGAGAGCATTACAGCCAACTTAAAGCCCAAGTTCTTACGAAT 242 | Db | 243 CCAAGTCTCTCCCTTCTATACTTTCGTAAGCCGATTCATCTTAACAAAGTAATGCCAACA 302 |
| Qy | 133 SerSerSerPheProSerProSerArgGlyGluProAsnAsnAsnMetSerSerThrPhe 152 | Qy | 133 SerSerSerPheProSerProSerArgGlyGluProAsnAsnAsnMetSerSerThrPhe 152 |
| Db | 303 TCCCGCTTCTCCCAAGTCATCTCGAGTGATGCT-----AACAGTGTCTGCATATCCA 356 | Db | 303 TCCCGCTTCTCCCAAGTCATCTCGAGTGATGCT-----AACAGTGTCTGCATATCCA 356 |
| Qy | 153 PheProPheLeuArgAsnGlyIleProSerSerLeuProSerLeuArgIleSerAsn 172 | Qy | 153 PheProPheLeuArgAsnGlyIleProSerSerLeuProSerLeuArgIleSerAsn 172 |
| Db | 357 TTGGCATTTCTC---CAATGTTCCATTTCCCTTGGGTACACACATTACGAATATCAAAAC 413 | Db | 357 TTGGCATTTCTC---CAATGTTCCATTTCCCTTGGGTACACACATTACGAATATCAAAAC 413 |
| Qy | 173 SerCysProValThrProProValSerSerProThrSerLysAsnProLysProLeuPro 192 | Qy | 173 SerCysProValThrProProValSerSerProThrSerLysAsnProLysProLeuPro 192 |
| Db | 414 AGTCCCTCTTAACACACACCTTTCTATCCAACT---AGAGTCCCTTAACACACATATT 470 | Db | 414 AGTCCCTCTTAACACACACCTTTCTATCCAACT---AGAGTCCCTTAACACACATATT 470 |
| Qy | 193 AsnTrpLeuSerIleAlaLysGlnSerMetAlaIleAlaLysGlnSerMetLysSerPhe 212 | Qy | 193 AsnTrpLeuSerIleAlaLysGlnSerMetAlaIleAlaLysGlnSerMetLysSerPhe 212 |
| Db | 471 AATCTTAGACT-----TTGGCTAGAGAGTCTATGTCGCTCTA 509 | Db | 471 AATCTTAGACT-----TTGGCTAGAGAGTCTATGTCGCTCTA 509 |
| Qy | 213 AsnTyrProPheTyrAlaValSerAlaProAlaSerProThrHisArgHisGlnPheHis 232 | Qy | 213 AsnTyrProPheTyrAlaValSerAlaProAlaSerProThrHisArgHisGlnPheHis 232 |
| Db | 510 AATATCCCTTTCTTTGGCTTCACGCCCACTATG3CCCACTCGAGGTCAACGATTC--- 566 | Db | 510 AATATCCCTTTCTTTGGCTTCACGCCCACTATG3CCCACTCGAGGTCAACGATTC--- 566 |
| Qy | 233 ThrLeuAlaThrIleProGluCysAspGlnSerAspSerSerThrValAspSerGlyHis 252 | Qy | 233 ThrLeuAlaThrIleProGluCysAspGlnSerAspSerSerThrValAspSerGlyHis 252 |
| Db | 567 ACTCTCTCAATAACAGAGGTGAGCAATCTATTCATCCACCATTTGATTTGGCCAG 626 | Db | 567 ACTCTCTCAATAACAGAGGTGAGCAATCTATTCATCCACCATTTGATTTGGCCAG 626 |
| Qy | 253 TrpIleSerPheGlnLysPheAlaGlnGlnGlnProPheSerAlaSerMetValProThr 272 | Qy | 253 TrpIleSerPheGlnLysPheAlaGlnGlnGlnProPheSerAlaSerMetValProThr 272 |
| Db | 627 TCGATGAGCTTCAAAGTAGCA-----GCCAATGGATCCCTACT 668 | Db | 627 TCGATGAGCTTCAAAGTAGCA-----GCCAATGGATCCCTACT 668 |
| Qy | 273 SerProThrPheAsnLeuValLysProAlaProGlnGlnMetSerProAsnThrAlaAla 292 | Qy | 273 SerProThrPheAsnLeuValLysProAlaProGlnGlnMetSerProAsnThrAlaAla 292 |
| Db | 669 TCTCCGACTTTAATCTTAATTAACCTGCTAGAGCTCAGAGAAATTCCTTCAATGATATGATC 728 | Db | 669 TCTCCGACTTTAATCTTAATTAACCTGCTAGAGCTCAGAGAAATTCCTTCAATGATATGATC 728 |
| Qy | 293 PheGlnGluIleGlyGlnSerSerGluPheLysPheGlu 305 | Qy | 293 PheGlnGluIleGlyGlnSerSerGluPheLysPheGlu 305 |
| Db | 729 ATCGACAAG--GGTAAGAGCATTTGAATTGACTTTGAG 764 | Db | 729 ATCGACAAG--GGTAAGAGCATTTGAATTGACTTTGAG 764 |
| RESULT 5 | 86887699 | RESULT 5 | 86887699 |
| LOCUS | BS887699 | LOCUS | BS887699 |
| DEFINITION | EST151550 cSTD Solanum tuberosum cDNA clone cSTD610 5' sequence, mRNA sequence. | DEFINITION | EST151550 cSTD Solanum tuberosum cDNA clone cSTD610 5' sequence, mRNA sequence. |

ACCESSION BG887699
 VERSION BG887699.1 GI:14264785
 KEYWORDS EST
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 768)
 van der Hoeven, R., Bezzerides, J., Ewing, E., Cho, J., Chimingo, A.,
 Bougri, O., Buell, C. R., Romling, C., Tanksley, S. and Baker, B.
 Generations of ESTs from dormant potato tubers
 Unpublished
 TITLE Contact: Robin Buell
 JOURNAL The Institute for Genomic Research
 COMMENT 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics
 Institute. Orders can be made through URL:
 http://genome.arizona.edu/orders/
 Seq primer: M13P-R.
 FEATURES
 source
 1..768
 location/Qualifiers
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cST6d10"
 /rname_type="dormant tuber"
 /dev_stage="one month post-harvest"
 /lab_host="SOLR"
 /clone_lib="cSTP"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; This library targets genes expressed in dormant
 tubers. This library was made from sections of dormant
 tuber, avoiding the buds and epidermis. Tubers were stored
 for one month post-harvest at 4°C. The tuber was peeled,
 well away from the surface. Then it was chopped into 1-2
 mm cubes and immediately frozen in liquid nitrogen. This
 library is noted as P4 in Tanksley lab notebooks."
 BASE COUNT 220 a 209 c 133 g 206 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.53e-55 Length: 768
 Score: 871.00 Matches: 183
 Percent Similarity: 77.07% Conservatives: 22
 Best Local Similarity: 68.80% Mismatches: 41
 Query Match: 49.35% Indels: 21
 DB: 12 Gaps: 6
 US-09-995-938a-7 (1-336) x BG887699 (1-768)
 QY 25 ProSerTrpArgGluArgGluAsnAsnArgArgArgGluArgArgArgAlaValAla 44
 DB 2 CCATCAGGAGAGAAAGAAATACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 61
 QY 45 AlAlbAlletTrpArgGluArgGluAlaGlnGlnAlaGlnGlnAlaGlnGlnAlaGlnGln 64
 DB 62 GCTTAGATTATATCTGTTTAAAGCTCTTAAAGCTCTTAAAGCTCTTAAAGCTCTTAAAGCT 121
 QY 65 AsnAsnGluValLeuLysAlaLeuCyValGluAlaGlyTTPValValGluGluAspGly 84
 DB 122 AACATGAAAGTTCTTAAAGCTCTTAAAGCTCTTAAAGCTCTTAAAGCTCTTAAAGCT 181
 QY 85 ThrThrTyrArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 104
 DB 182 ACCACTATATGCAAGGATGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 241
 QY 105 ValThrProTyrSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 124
 DB 242 ATCAGCGCAAGTTCTTAAAGCTCTTAAAGCTCTTAAAGCTCTTAAAGCTCTTAAAGCT 301

QY 125 ProSerTrpGlnValSerProSerSerSerSerSerSerSerSerSerSerSerSerSerSer 144
 DB 302 CCATCTTATACAGCAAGTCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 361
 QY 145 AsnAsnAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 164
 DB 362 -----AACATGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 412
 QY 165 LeuProSerLeuArgGlnSerSerSerSerSerSerSerSerSerSerSerSerSerSer 184
 DB 413 CTTCCTCATATGCAAGTAAATACAGAGTCCCTCTTAACTCCACCTCTTCACTCACTCACT 472
 QY 185 SerLysAsnProLysProLysProLysProLysProLysProLysProLysProLysPro 204
 DB 473 ---AGCATCTTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 508
 QY 205 AlAlaGlnSerMetAlaSerAlaSerAlaSerAlaSerAlaSerAlaSerAlaSerAlaSer 224
 DB 509 GCCAAGATCAATCATGTTGCTTAAACATCCCTCTTCTGCTGCTGCTGCTGCTGCTGCTG 568
 QY 225 ProThrHisArgHisGlnPheHisThrLeuAlaThrIleProGluCyAspGluSerAs 244
 DB 569 CCAACTAGAGTACGCTTTT---ACTCTCTCCCAACTATACCCAGTGTGATGAATCTGA 625
 QY 244 pSerSerThrValAspSerGlyHisTrpIleSerPheGlnLysPheAlaGlnGlnLys 264
 DB 626 CTATCTACATCATATATTCAGCCAGTATATCACTTCAAAAGTAC----- 671
 QY 264 oPheSerAlaSerMetValProThrSerProThrPheAsnLeuValLysProAlaProG 284
 DB 672 -----GCCGCAAGCTTCCACCTTCTCCACA-TTAAATTTTAAACCTGTGCTCA 723
 QY 284 ngLmeSerProAsn 289
 DB 724 GCCACTTCGTCTTAAT 739
 RESULT 6
 AV551486 585 bp mRNA linear EST 06-SEP-2000
 LOCUS AV551486 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
 DEFINITION cDNA clone R2127d10R 5', mRNA sequence.
 AV551486
 ACCESSION AV551486.1 GI:8722899
 VERSION AV551486.1
 KEYWORDS EST
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophytes;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 / euroside II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 585)
 Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized end
 size-selected cDNA libraries
 DNA Res. 7, 175-180 (2000)
 JOURNAL MEDLINE 20363093
 PUBMED 10907847
 COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 FEATURES
 source
 1..585
 location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="R2127d10R"
 /rname_type="roots"
 /clone_lib="Arabidopsis thaliana roots Columbia"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"

| Accession | Source | Organism | Keywords | Version | Accession | Source | Organism | Keywords | Version |
|-----------|--|--|--|--|-----------|---|-------------------|----------|-------------|
| 566 | CGTGTACAAATACAGAGTGTGATGAGCTGCTATTCATCCACCATGATTTGGCCAGTG | Solanum tuberosum | EST | 07-MAR-2003 | 254 | 115erPheginLyaPheAlGInGInLnProPheSerAlaser-MetvalProThrS | Solanum tuberosum | EST | 07-MAR-2003 |
| 625 | ATGAGCTTTCAAAAGTATGCA-----GCCAATGGGAGTCCCTACTTC | Solanum tuberosum | EST | 07-MAR-2003 | 273 | ProThrPheMetLeuValLysProAlaProGInGInMetSerProAntThrAlaAlaP | Solanum tuberosum | EST | 07-MAR-2003 |
| 667 | 626 ATGAGCTTTCAAAAGTATGCA-----GCCAATGGGAGTCCCTACTTC | Solanum tuberosum | EST | 07-MAR-2003 | 293 | 668 TCCGACTTTTATCTTATTAAAGCCGTAGCTCAGAGATTTCTCTTATATGATGATCAT | Solanum tuberosum | EST | 07-MAR-2003 |
| 727 | 273 ProThrPheMetLeuValLysProAlaProGInGInMetSerProAntThrAlaAlaP | Solanum tuberosum | EST | 07-MAR-2003 | 728 | CGACACAG---GCTAAGAGCATTTGAGCTTGACTTTGAG | Solanum tuberosum | EST | 07-MAR-2003 |
| 761 | 293 eGInGInLLeGlyGInSerSerGlyLysPheLysPheGlu | Solanum tuberosum | EST | 07-MAR-2003 | | | | | |
| | 766 bp | | | | | | | | |
| | EST492505 | EST492505 | EST492505 | EST492505 | | | | | |
| | CGS53827 | CGS53827 | CGS53827 | CGS53827 | | | | | |
| | GI:13611967 | GI:13611967 | GI:13611967 | GI:13611967 | | | | | |
| | EST | EST | EST | EST | | | | | |
| | Solanum tuberosum (potato) | Solanum tuberosum | Solanum tuberosum | Solanum tuberosum | | | | | |
| | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | | | | |
| | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | | | | |
| | Asteridae; Lamiales; Solanales; Solanaceae; Solanum. | Asteridae; Lamiales; Solanales; Solanaceae; Solanum. | Asteridae; Lamiales; Solanales; Solanaceae; Solanum. | Asteridae; Lamiales; Solanales; Solanaceae; Solanum. | | | | | |
| | 1 (baas 1 to 766) | 1 (baas 1 to 766) | 1 (baas 1 to 766) | 1 (baas 1 to 766) | | | | | |
| | van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemingo, A., | van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemingo, A., | van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemingo, A., | van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemingo, A., | | | | | |
| | Bougrit, O., Buell, C.R., Romling, C., Tanksey, S. and Baker, B. | Bougrit, O., Buell, C.R., Romling, C., Tanksey, S. and Baker, B. | Bougrit, O., Buell, C.R., Romling, C., Tanksey, S. and Baker, B. | Bougrit, O., Buell, C.R., Romling, C., Tanksey, S. and Baker, B. | | | | | |
| | Generations of ESTs from sprouting potato eyes | Generations of ESTs from sprouting potato eyes | Generations of ESTs from sprouting potato eyes | Generations of ESTs from sprouting potato eyes | | | | | |
| | Unpublished | Unpublished | Unpublished | Unpublished | | | | | |
| | Contact: Robin Buell | Contact: Robin Buell | Contact: Robin Buell | Contact: Robin Buell | | | | | |
| | The Institute for Genomic Research | The Institute for Genomic Research | The Institute for Genomic Research | The Institute for Genomic Research | | | | | |
| | 9712 Medical Center Dr, Rockville, MD 20850, USA | 9712 Medical Center Dr, Rockville, MD 20850, USA | 9712 Medical Center Dr, Rockville, MD 20850, USA | 9712 Medical Center Dr, Rockville, MD 20850, USA | | | | | |
| | Email: potato-array@igr.org | Email: potato-array@igr.org | Email: potato-array@igr.org | Email: potato-array@igr.org | | | | | |
| | This clone can be obtained from the University of Arizona Genomics | This clone can be obtained from the University of Arizona Genomics | This clone can be obtained from the University of Arizona Genomics | This clone can be obtained from the University of Arizona Genomics | | | | | |
| | Institute. Orders can be made through URL: | Institute. Orders can be made through URL: | Institute. Orders can be made through URL: | Institute. Orders can be made through URL: | | | | | |
| | http://genome.arizona.edu/orders/ | http://genome.arizona.edu/orders/ | http://genome.arizona.edu/orders/ | http://genome.arizona.edu/orders/ | | | | | |
| | Seq primer: M13F-R. | Seq primer: M13F-R. | Seq primer: M13F-R. | Seq primer: M13F-R. | | | | | |
| | Location/Qualifiers | Location/Qualifiers | Location/Qualifiers | Location/Qualifiers | | | | | |
| | 1..766 | 1..766 | 1..766 | 1..766 | | | | | |
| | /organism="Solanum tuberosum" | /organism="Solanum tuberosum" | /organism="Solanum tuberosum" | /organism="Solanum tuberosum" | | | | | |
| | /mol_type="cDNA" | /mol_type="cDNA" | /mol_type="cDNA" | /mol_type="cDNA" | | | | | |
| | /cultivar="Kennebec" | /cultivar="Kennebec" | /cultivar="Kennebec" | /cultivar="Kennebec" | | | | | |
| | /db_xref="taxon:4113" | /db_xref="taxon:4113" | /db_xref="taxon:4113" | /db_xref="taxon:4113" | | | | | |
| | /clone="CGS53827" | /clone="CGS53827" | /clone="CGS53827" | /clone="CGS53827" | | | | | |
| | /feature_type="sprouting eyes from tubers" | /feature_type="sprouting eyes from tubers" | /feature_type="sprouting eyes from tubers" | /feature_type="sprouting eyes from tubers" | | | | | |
| | /dev_stage="12-14 weeks post harvest" | /dev_stage="12-14 weeks post harvest" | /dev_stage="12-14 weeks post harvest" | /dev_stage="12-14 weeks post harvest" | | | | | |
| | /lab_host="SOLR" | /lab_host="SOLR" | /lab_host="SOLR" | /lab_host="SOLR" | | | | | |
| | /clone_1lb="cDNA" | /clone_1lb="cDNA" | /clone_1lb="cDNA | | | | | | |

| Db | Accession | Source | Organism |
|------------|--|--|-----------------|
| Db | GAAGAAGCAATTC-AGGAAAGAGAGAGAAAGACGAGACTGTAGCTGCTAAGATTAT | | 61 |
| Qy | 49 Th-GlyLeuArgAlaGlnGlyAspTyrAsnLeuProLysHisCyAspAsnAsnGluVal | | 68 |
| Db | 62 ACTGGTTTAAGAGCTCAAGGAAACCTAAATCTTCTTAAACACTGTGATTAACAAAGAATT | | 121 |
| Qy | 69 LeuValAlaLeuCyAsnValGluAlaGlyTyrValValGluGluAspGlyThrThrTyrArg | | 88 |
| Db | 122 CTTAAAGCTCTTGTACTGAAGCTGGTGGATCGTGAACCTGATGGTACCATTATTCGC | | 181 |
| Qy | 89 LysGlyCyAspLeuProGluGluLeuAlaGlyThrSerSerArgValThrProTyr | | 108 |
| Db | 182 AAGGAGTCAAGCCACCCGATGGAGATGTGAGGCACTTCAACAAACATCAGCCCAAGT | | 241 |
| Qy | 109 SerSerGlnAsnGlnSerProLeuSerSerAlaPheGlnSerProIleProSerTyrGln | | 128 |
| Db | 242 TCTTCAAGCAGCATCAAGTCCCGCTGATCACTTGTCTAGCCCAATTCATCTTATCAG | | 301 |
| Qy | 129 ValSerProSerSerSerSerPheProSerProSerArgGluProAsnAsnMet | | 148 |
| Db | 302 CCAAGTCAACCTCTCTTCTTCCCAAGCCCACTCCGTCTGATGCT---AACATG | | 355 |
| Qy | 149 SerSerThrPhePheProPheLeuArgAsnGlyGlyIleProSerSerLeuProSerLeu | | 168 |
| Db | 356 TCATCATCATCATATTTCTTCTCCATTA---GTCACTCCCTCATCTCTTCCATCATTA | | 412 |
| Qy | 169 ArgIleSerAsnSerCysProValThrProProValSerSerProThrSerLeuAsnPro | | 188 |
| Db | 413 CGAATATCAAAAGTGCCCTGTAACTCCACCTCTTCAATCACCACACT---AGGCATCT | | 469 |
| Qy | 189 LysProLeuProAsnTyrGluSerIleAlaLysGlnSerMetAlaIleAlaLysGlnSer | | 208 |
| Db | 470 AAGCAAACTTCAATTAAAGAC-----TTGGCCAAAGAAATCA | | 508 |
| Qy | 209 MetAlaSerPheAsnTyrProPheTyrAlaValSerAlaProAlaSerProThrHisArg | | 228 |
| Db | 509 ATGTTTCCTTAAACATCCCTTCTTCTTCTGCTTCAAGCCCAAGCAACCTAGGGGT | | 568 |
| Qy | 229 HisGlnPheHisThrLeuAlaThrIleProGluCysAspGluSerAspSerSerThrVal | | 248 |
| Db | 569 CAGCGTTT----- | | 577 |
| Qy | 249 AspSerGlyHisTyrPheSerPheGlnLysPheAlaGlnGlnGlnProPheSerAlaSer | | 268 |
| Db | 577 ----- | | 577 |
| Qy | 269 MetValProThrSerProThrPheAsnLeuValLysProAlaProGlnGlnMetSerPro | | 288 |
| Db | 578 -----ACTCTCCAACTTTAATCTTGTAAACCTGCTCAGCCACTTGTCTCT | | 628 |
| Qy | 289 AsnThrAlaAlaPheGlnGluIleGlyGlnSerSerGluPheLysPheGluAsnSerGln | | 308 |
| Db | 629 AATAT---ATGATCACAGACAAAGGTAAGCATAGACTTCGACTTGAATAATGATCA | | 685 |
| Qy | 309 ValLysProTyrGluGlyGluArgIleHisAspValGluMetGluAspLeuGluLeuThr | | 328 |
| Db | 686 GTCAAGGCATCGGAAAGGGAAGGATTCAGAAATGTAATCGATGATCGGACTGACACA | | 745 |
| Qy | 329 LeuGlyAsnGlyLysAlaArg | | 335 |
| Db | 746 CTTGGAAGTGCAATGCTCTGC | | 766 |
| RESULT 9 | BGI30895 | 658 bp | mRNA |
| LOCUS | EST463787 | tomato crown gall Lycopersicon esculentum cDNA clone | EST 31-JAN-2001 |
| DEFINITION | CTDIE186.5 | sequence, mRNA | sequence. |
| ACCESSION | BGI30895 | | |
| VERSION | BGI30895.1 | | GI:12631083 |
| KEYWORDS | EST. | | |
| SOURCE | Lycopersicon esculentum (tomato) | | |
| ORGANISM | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |

Db 122 GGGGAT-----AACATGCGCTTTTAATCTCCTTCGATCTTCGAGT---GCCATT 172
 Qy 162 ProSerSerLeuProSerLeuArgIleSerAaspSerProValThrProProValSer 181
 Db 173 CCATGTCCTCTCTCTCTCTCGAATCTCAACAGTCGCGCTTCAACCCGCTCTCTCG 232
 Qy 182 SerProHisSerLeuAaspProIleProleuProantProIleuAlaIleValGlnSer 201
 Db 233 TCCCGGCTCCCAAGAAACCCCAACCAATTCCTCAACTGGGAC----- 274
 Qy 202 MetAlaIleAlaValGlnSerMetAlaSerPheAntYrProPheYrAlaValSerAla 221
 Db 275 ---TTTATTGCCAAACATCATCGCTCTCTTATTACCATTTAAGCGGTCTGCT 331
 Qy 222 ProAlaSerProThrHisArgHisGlnPheHisThrLeuAlaThrIleProGluCysAap 241
 Db 332 CCGGCTAGCCCACTCAACCGT---CAGTTTCATCTCTCAGCAGCATATACGTAATGTGAT 388
 Qy 242 GluSerAaspSerThrValAaspSerGlyHisThrIleSerPheGlnLysPheAlaGln 261
 Db 389 GAGCTGATATCATCTCACTGTGAGTCTGTGTCAGTGAATACCTTTCAAAAGTTTGGC-- 445
 Qy 262 GlnGlnProPheSerAlaSerMetValProThrSerProThrPheAaspValIlePro 281
 Db 446 -----CTTCTGTGCTGCGACGATGCCAACCCTCTCTACTTAATCTTGTGATACCC 499
 Qy 282 AlaProGlnIleMetSerProAaspThrAlaAlaPheGlnIleGlyGlnSerSerGlu 301
 Db 500 GTGGCTCAGCAAAATTGCTCAACAATTGGTCAAGAGAAATGA---GTGCCCATGGAT 556
 Qy 302 PheLysPheGlnLysAaspSerGlnValIleProThrIleGluIleHisAaspValIle 321
 Db 557 TTTCAGATTGTTGTTGATGAACGGTGAAACCATGGAGAGAGAGAGATTCAGAGTAGGA 616
 Qy 322 MetGluAaspLeuGlnLeuThrLeuGly 330
 Db 617 TTAGATGATCTAGAGCTCACCTTGA 643

RESULT 11
 AV440980/c 619 bp mRNA linear EST 14-NOV-2000
 LOCUS AV440980 Arabidopsis thaliana above-ground organ two to six-week
 DEFINITION Old Arabidopsis thaliana cDNA clone AP218b09_f3, mRNA sequence.
 ACCESSION AV440980
 VERSION AV440980.1 GI:7611357
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (chale crese)
 ORGANISM Arabidopsis thaliana
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 619)
 REFERENCE
 AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 JOURNAL DNA Res. 7, 175-180 (2000)
 MEDLINE 20363093
 PUBMED 10907847
 COMMENT
 Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yena 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 Location/Qualifiers
 1. 619
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="AP218b09_f"
 /tissue_type="aboveground organs"
 /dev_stage="two to six-week old"

/clone_1b="Arabidopsis thaliana above-ground organ two to
 six-week old"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"
 BASE COUNT 176 a 129 c 148 g 166 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,41e-48 Length: 619
 Score: 776.50 Matches: 147
 Percent Similarity: 92.07% Conservative: 4
 Best Local Similarity: 89.63% Mismatches: 12
 Query Match: 43.99% Indels: 1
 DB: 9 Gaps: 1
 US-09-995-938a-7 (1-336) x AV440980 (1-619)

Qy 172 AaspSerCysProValThrProProValIleSerSerProThrSerLysAaspProIleu 191
 Db 618 AACAGTCTCTCTGTCATCTCCACGATGTCATCCCACTTACGAAACCCAAACCATTTG 559
 Qy 192 ProAaspThrIleSerLeuAlaValGlnSerMetAlaIle--AlaValGlnSerMetAla 210
 Db 558 CCAATCTGGGAATCTTTACCAACAAATCATGTCATGCTGCTTAAACAGTCAATGACT 499
 Qy 211 SerPheAaspYrProPheYrAlaValSerAlaProAlaSerProThrHisArgHisGln 230
 Db 498 TCTTTGAATCAACCGTTTATGCGGTGTGTCACCTGCACTCTCACTCATCATCGCAG 439
 Qy 231 PheHisThrLeuAlaThrIleProGluCysAaspIleSerAaspSerThrValAaspSer 250
 Db 438 TTCATGCTCCGGTACTAATCACTGAATGTAGTGAAGTCACTTCCACTGTATATCT 379
 Qy 251 GlyHisThrIleSerPheGlnLysPheAlaGlnIleGlnIleProPheSerAlaSerMetVal 270
 Db 378 GGTGATGATGATGATCTTCAAAAGTTTGCACAAACAGCCATCTCGCTATGTGG 319
 Qy 271 ProThrSerProThrPheAaspLeuValIleProAlaProGlnIleMetSerProAaspThr 290
 Db 318 CCAACTCGCCCTACCTTCAATCTCGTGAACCTGCACACAGCAATTTCTTCCAAACACA 259
 Qy 291 AlaAlaPheGlnIleGlyGlnSerSerGluPheLysPheGlnLysAaspSerGlnValIle 310
 Db 258 GCAGCAATCCAAAGATTTGGCAAGCTCCGAGTTTGAAGTGAAGACGCAACATTAAG 199
 Qy 311 ProThrIleGluIleGlyGlnIleHisAaspValGlyMetGluAaspLeuGlnLeuThrLeuGly 330
 Db 198 CCATGGGAAGGGAGAGATCCATGATGTGCTATGGAGATCTAGAGCTCACGCTTGA 139
 Qy 331 AaspGlyLysAla 334
 Db 138 AATGTTAAAGCT 127

RESULT 12
 B12579 941 bp DNA linear GSS 14-MAY-1997
 LOCUS F22H5-T7 IGF Arabidopsis thaliana genomic clone F22H5, genomic
 DEFINITION survey sequence.
 ACCESSION B12579
 VERSION B12579.1 GI:2093699
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (chale crese)
 ORGANISM Arabidopsis thaliana
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 941)
 REFERENCE
 AUTHORS Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shim, P., Sun, H. and
 Ecker, J.
 TITLE BAC End Sequences at ATGC
 JOURNAL Unpublished
 COMMENT Other GSSs: F22H5-SP6
 Contact: Ecker J.

US-09-995-938A-7 (1-336) x US-09-313-294A-6581 (1-196)

OY 50 GYLeuAARGAlaGlnGlyAspTyrAenLeuProLysHisCysAspAsnGlnValLeu 69
DB 5 GGCCTCCGGCCGACGGCGCTACAGCTGCCCAAGCTGCGACAAAC----- 55
OY 70 LySAlaLeuCyValAlaGlnGlyTTPValValGlnGlyAspGlyThrThrTyrArgLys 89
DB 56 -----GAGGCGGCTGGATCTCGAGCCCGACGACCACTACCGCCAG 100
OY 90 Gly 90
DB 101 GGA 103
RESULT 2
US-09-249-585A-4
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/00905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4
Alignment Scores:
Pred. No.: 0.00251 Length: 1926
Score: 143.00 Matches: 70
Percent Similarity: 40.89% Conservative: 31
Best Local Similarity: 28.34% Mismatches: 100
Query Match: 8.10% Indels: 47
DB: 4 Gaps: 8
US-09-995-938A-7 (1-336) x US-09-249-585A-4 (1-1926)
OY 95 ProGlyGluIleAlaGlyThrSerSerArgValThrProTyrSer---SergInaGln 113
DB 540 CCGTCTCCCGCCGCTCCGCTCTCTCCGCTCCCGCTCCCGCTCCCGCTCCGCTCC 599
OY 114 SerProLeuSerSerAlaPheGlnSerProIleProSerTyrGlnValSerProSer 133
DB 600 TCCTCGTCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 659
OY 134 SerSerPhePro-----SerProSerArgGlyGluPro 144
DB 660 TCCTCTCCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 719
OY 145 AaAnaAaMeSerSerThrPhePheProPheLeuArgAsnGlyIlePro----- 162
DB 720 TCCTGCTCTCCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 779
OY 163 -----SerSerLeuProSerLeuArgIleSer 171
DB 780 CCGTCTCCGCTCCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 839
OY 172 AaSerCysProValThrProProValSerSerProThrSerIleAsnProLysProLeu 191
DB 840 TCCTCTCCCGCTCCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 896
OY 192 ProAenTyrGluSerIleAlaLysGlnSerMetAlaIleAlaLysGlnSerMetAlaSer 211
DB 897 CCGTCTCCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 956
OY 212 PheAenTyrProPheTyrAlaValSerAlaProAlaSerProThrIleArgHisGlnPhe 231

DB 957 TCCTCCCGCTCTCTCTCCACCTTCGCGCCCAAGCTCTCCGCTACCTCCGCGCCAG 1016
OY 232 HisThrLeuAlaThrIleProGluCysAspGlnSerAspSerSerThrValAspSergly 251
DB 1017 TCCTCATTACCTCCGCGCC-----AGCTCTCCATTACCTCCGCG 1058
OY 252 HisThrIleSerPheGlnLysPheAlaGlnGlnInProPheSerAlaSerMetValPro 271
DB 1059 GCGCCCATCTCTGCACT-----TTCTCGATCCCGCCCTTCAGCACTTTCGATCC 1112
OY 272 ThrSerProThrPheAsnLeuValysProAlaProGln-----GlnMet 286
DB 1113 CTCTCCAGACCTG-CACCTCTTTCTCCGGGCTCTCAGGTCATCAGTACGTAGTAGTA 1171
OY 287 SerProAenThrAlaAlaPheGlnGlnIleGlyGlnSerSerGluPheLysPheGlnAa 306
DB 1172 GCGCCAGAGTGGCGCCGCTCCGCGGAGATC-----CAT 1204
OY 307 SerGlnValLysProTyrGlu 313
DB 1205 CTTCGGTAAAGGTGGAC 1225
RESULT 3
US-09-130-114-2
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Robbins, Alan K.
; APPLICANT: Dams, Bassem B.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FILE REFERENCE: 0867/10903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2
Alignment Scores:
Pred. No.: 0.00252 Length: 1931
Score: 143.00 Matches: 70
Percent Similarity: 40.89% Conservative: 31
Best Local Similarity: 28.34% Mismatches: 100
Query Match: 8.10% Indels: 47
DB: 2 Gaps: 8
US-09-995-938A-7 (1-336) x US-09-130-114-2 (1-1911)
OY 95 ProGlyGluIleAlaGlyThrSerSerArgValThrProTyrSer---SergInaGln 113
DB 540 CCGTCTCCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 599
OY 114 SerProLeuSerSerAlaPheGlnSerProIleProSerTyrGlnValSerProSer 133
DB 600 TCCTCGTCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 659
OY 134 SerSerPhePro-----SerProSerArgGlyGluPro 144
DB 660 TCCTCTCCCGCTCTCCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 719
OY 145 AaAnaAaMeSerSerThrPhePheProPheLeuArgAsnGlyIlePro----- 162
DB 720 TCCTGCTCTCCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 779
OY 163 -----SerSerLeuProSerLeuArgIleSer 171
DB 780 CCGTCTCCGCTCCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 839

```
Gy      172 AenserrCySPProValThPrPProvAlSerSPrOTHrSeTylSaSnPrOlyAePrOUen 191  
       :|||||::|||:|||||::|||:|||||::|||:|||||::|||:  
Db      840 TCCTCCTCCCGGTCTGTCGCCTGCCGGTTCTTGCTCCGCTCCCGGTCTGCTCC--- 896  
  
Gy      192 ProAntnTpGuSErlEalElYgSlnerMetAlIleAlLyalYLglInserMeLaSer 211  
       :|||||::|||:|||||::|||:|||||::|||:|||||::|||:  
Db      897 CGGTCTCCCGGTCTGTCGCCTGCCGGTTCTTGCTCCGCTCCCGGTCTGCTCCCGGTCTCG 956  
  
Gy      212 PheAnTYrPRoPhNeTyALaVAlSERAlAProAlAserPROTHrhIArgHlgnDphe 231  
       ::|||::|||::|||::|||::|||::|||  
Db      957 TCTCTCCCGGTCTGCTCCGCTCCAGCTCCGACCCCACTCTCCGTCACTCCGGCCCCAGC 1016  
  
Gy      232 HIsThrILeuAlEthIrLePProgLuYaARguISerbseSThVaLAbseryl 251  
       :|||||::|||:|||||::|||:|||||::|||:|||||::|||:  
Db      1017 TCCTCATCACCTCCGGCCCC-----AGTCTTCATCACCTCCGGC 1058  
  
Gy      252 HisTPilesErPheGLnlYPheAlnglnInglnPrOpHeSeRAlASerMeVALpro 271  
       |||||::|||:|||||::|||:|||||::|||:|||||::|||:  
Db      1059 GGGCCATCTCTGCACT- -TTCGGGTCCCCTTCAGACTTTCGGTCCC 1112  
  
Gy      272 ThrsEPROTrHPheaSLnuvalLYsPrOAPlProgIn-----Glmet 286  
       ::|||::|||::|||::|||::|||::|||  
Db      1113 CTCTCCACCACCTCG-CACCTCTTTTTCTCGGGTCTCGAGGTCACTACGTCACTAGTAGTA 1171  
  
Gy      287 SerPROanTRHAalaPHeslngluIEglInsersergluPhLyAPheGUasn 306  
       |||||::|||:|||||::|||:|||||::|||:|||||::|||:  
Db      1172 GGCCcAgGrTgcCcCtCGggGaGgtC-----CAT 1204  
  
Gy      307 serGINvaLLysPROtpGLU 313  
       |||||::|||:|||||::|||:|||||::|||:|||||::|||:  
Db      1205 CTTCCGTAAAAAAGGTGGGAC 1225
```

RESULT 4

```
US-08-458-568A-11  
; Sequence 11, Application US/08458568A  
; Patent No. 5821339  
GENERAL INFORMATION:  
APPLICANT: Schaffer, Patricia A.  
APPLICANT: Yeh, Lily  
TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesviruses  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESS: Woodcock, Washburn, Kurtz, Mackiewicz & NO. 5821339.rts  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,568A  
FILING DATE: 02-JUNE-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/065,146  
FILING DATE: 05-MAY-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Leary Ph.D., Kathryn R.  
REGISTRATION NUMBER: 36,317  
REFERENCE /DOCKET NUMBER: DPCI-0029  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO.: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12001 base pairs
```

[illegible]


```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,441
FILING DATE: 20-OCT-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,170
FILING DATE: 13-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/010,409
FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ethnozn, Gregory F.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 015280-113110US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9636 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 149..9556
US-08-954-441-1

```

| | |
|------------------------|--------|
| Alignment Scores: | |
| pred. No.: | 0.0605 |
| Score: | 140.00 |
| Percent Similarity: | 37.60% |
| Best Local Similarity: | 26.10% |
| Query Match: | 7.93% |
| DB: | 4 |
| | 9 |
| Length: | 9636 |
| Matches: | 68 |
| Conservative: | 23 |
| Mismatches: | 78 |
| Indels: | 74 |
| Gaps: | 9 |

US-09-995-938A-7 (1-336) x US-08-954-441-1 (1-9636)

| | | | |
|----|------|--|------|
| Qy | 102 | SerSerArgValItnrProTyrSerSerGlnAsnGlnSerProLeuSerSerAlaIleGln | 121 |
| Db | 1488 | TCATCTACATATATACCTCTTCTTTCAGCTACCTACCTTCCTACCTCTTCTTTCAGCTACCTCT | 1429 |
| Qy | 122 | SerProIleProSerTyrGlnValSerProSerSerSerSerPheProSerProSerArg | 141 |
| Db | 1428 | TCACCTACCCCTTTCACCTACCTCTTTCACCTCTTCTTTCACCTACCTCTTTCACCTCT | 1369 |
| Qy | 142 | GlyGlnProAsnAsnAsnMetSerSerThrPhePheProPheLeuArgAsnGlyIle | 161 |
| Db | 1368 | TCACCTACCTCTTTCACCTACCTCTTCA----- | 1342 |
| Qy | 162 | ProSerSerLeuProSerLeuArgIleSerAsnSerCyProValThrProProValSer | 181 |
| Db | 1341 | CTTACTCTTTCACCTACCTACCT-----TCTTCACTACCTCTTTCACCTACCTCT | 1297 |
| Qy | 182 | SerProThrSerTyrAsnProIleProLeu-----ProAsn | 193 |
| Db | 1296 | TCACCTACCTCTTTCATCATCATCTTCATTTTATTTATGATCATCATTAAGTATACCACTCT | 1237 |
| Qy | 194 | TrpGlnSer-----IleAlaIysGlnSerMetAlaIleAlaIysGln | 207 |
| Db | 1236 | TGGTCTTATATATATTCGTCGAACCTTTATCATATGCTCATCCGCTCTTTCACAGAACT | 1177 |
| Qy | 208 | SerMetAlaSerPheAsnTyrProPheTyrAlaValSerAlaProIleSer-Pro----- | 225 |
| Db | 1176 | TCATCCTCATCATTAATAACTTCT--TCTGAAGATTCAGCAGCACTTCTTCGATGT | 1120 |
| Qy | 226 | -----ThrHisArgHisGlnPheHisIle | 233 |
| Db | 1119 | TCTCTTGTATATGTTCTTCTGCTTTTCATGCCACTTCTCTTCAATAACATATATCATCATAT | 1060 |

[illegible]

RESULT 7
 US-08-224-482-3
 ; Sequence 3, Application US/08224482
 ; Patent No. 5837692
 ; GENERAL INFORMATION:
 ; APPLICANT: Mercola, Dan
 ; APPLICANT: Adamsom, Eileen D.
 ; TITLE OF INVENTION: Inhibition of the Mitogenic Activity of
 ; TITLE OF INVENTION: PCGF by Mammalian EGR
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:

ADDRESS: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA

COMPUTER READABLE;

COMPUTER: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/224,482

CLASSIFICATION: 435

NAME: Campbell, Cathryn A.

REFERENCE/DOCKET NUMBER: P-ME 9913

TELEPHONE: (619) 535-9001

; INFORMATION FOR SEQ ID NO: 3:

LENGTH: 3132 base pairs

STRANDEDNESS: do

US-08-224-482-3

Alignment Scores

Score: _____

Best Local Similarity

DB:

US-09-995-938A-7 (1-336) x US-08-224-482-3 (1-3132

```

QY      2 ThrSerAspGlyAlaThrSerThrSer-----AlaAlaAlaAlaAla 15

```

```

Db      1328  ACGAGCTCAACCCGACATCCGATCCACAGAGCCGACGAAAGCCCTTCAGTGGCCATCT 1387
Oy      16  AAlAAlAAlAAlAAlA-ArGArGlybProSerTPrArGluArGluAaAaAaArGAr 35
Db      1388  GGAATGGCAACTTCACGCGGACGACCACTTCACCAACCCACATCCGACCCACACAGAGCG 1447
Oy      35  gArGluArGArGArGArGAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAl 55
Db      1448  AAAAGCCCTTCGCTGCGACATCTGTGGAAGAAAGTTGCCAGAGAGGATGAACGACAAAG 1507
Oy      55  yAePTyRAnLeuProLybHisCyAaPAnAaAngluAlAAlAAlAAlAAlAAlAAlAAl 75
Db      1508  GGCATACCAAGATCCACTTCGCGGACAGAAAGCAAGAAAGCAAAAGTGTGTT----- 1562
Oy      75  uAlAglYTrPvAlValGlulAaPglY-ThrThyTrArGlybGlyCybLybProLeup 95
Db      1563  -----GGCCCTTCGCGGACATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1597
Oy      95  roGlyGlulAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlA 115
Db      1598  CGTCCCGGTTGCT-----ACCTCTTACCCGCTCCCGGTTACTACTACTCTTATCCATCCC 1651
Oy      115  rOlEusErSerAlAheGlnSerProIlEProSerTyRGlNValSerProSerSers 135
Db      1652  CGGCGCACACCTCATACCCATCCCTGTGCGCACCTCTCTCTCTCTCTCTCTCTCTCTCTCA 1711
Oy      135  eRPhEProSerPro--SerArGlyGlupProAaAaAaAaAaAaAaAaAaAaAaAaAaAa 154
Db      1712  CTTACCATCCCTCTGTGACAGAGGCTCCCTCCCTCCGTCGTCGACACACATAC----- 1767
Oy      154  rOPheLeuArGAnGlyGlYlePro-----SerSerLeuProS 167
Db      1768  -----TCCTCTGTTCCTCCCTGCTTCCTCCCGGCGGACGATGAGTTCCTT 1813
Oy      167  eRLeuArGlyLeSerAnSerCybProValThrProProValSerSerProThSer---- 185
Db      1814  CTTAGCTGTACCACTCTCTTACAGGCGCTCCACAGGCGCTTTCGACATGACAGCAACT 1873
Oy      185  ----- 185
Db      1874  TTTCTCCAGCAATTGAATTGCTTAAAGGAAAGGAAAGGAAAGGAAAGGAGAGA 1933
Oy      186  --LyAsnProLybProLeuProAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 205
Db      1934  AAAAGAAACAAAGACATTAAAGACAGAG-----GAGGAGATGGCCATAG 1981
Oy      205  lAlYglnSerMetAlaSerPheAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 225
Db      1982  GAGAGAGGTTCTCTTAGTCAGATGAGGTTCT-----TCAGAGCAAGTCTTC 2032
Oy      225  rOThr 226
Db      2033  CCTCT 2037

```

RESULT 8

```

US-09-205-921-1
; Sequence 1, Application US/09205921A
; Patent No. 6008048
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: ex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF EGR-1 EXPRESSION
; FILE REFERENCE: RTS-0028
; CURRENT APPLICATION NUMBER: US/09/205,921A
; CURRENT FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 3132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (271)..(1902)

```

US-09-995-938a-7 (1-336) x US-09-205-921-1 (1-3132)

```

Alignment Scores:
Pred. No.: 0,0171 Length: 3132
Score: 187,50 Matches: 63
Percent Similarity: 43,064 Conservative: 45
Best Local Similarity: 28,954 Mismatches: 92
Query Match: 7,794 Indels: 63
DB: 3 Gaps: 9

```

```

US-09-995-938a-7 (1-336) x US-09-205-921-1 (1-3132)
Oy      2  ThrSerbGlyAlaTherThSer-----AlaAlAAlAAlA 15
Db      1328  ACGAGCTCAACCCGACATCCGATCCACAGAGCCGACGAAAGCCCTTCAGTGGCCATCT 1387
Oy      16  AAlAAlAAlAAlAAlA-ArGArGlybProSerTPrArGluArGluAaAaAaArGAr 35
Db      1388  GGAATGGCAACTTCACGCGGACGACCACTTCACCAACCCACATCCGACCCACACAGAGCG 1447
Oy      35  gArGluArGArGArGArGAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAl 55
Db      1448  AAAAGCCCTTCGCTGCGACATCTGTGGAAGAAAGTTGCCAGAGAGGATGAACGACAAAG 1507
Oy      55  yAePTyRAnLeuProLybHisCyAaPAnAaAngluAlAAlAAlAAlAAlAAlAAlAAl 75
Db      1508  GGCATACCAAGATCCACTTCGCGGACAGAAAGCAAGAAAGCAAAAGTGTGTT----- 1562
Oy      75  uAlAglYTrPvAlValGlulAaPglY-ThrThyTrArGlybGlyCybLybProLeup 95
Db      1563  -----GGCCCTTCGCGGACATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1597
Oy      95  roGlyGlulAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlAAlA 115
Db      1598  CGTCCCGGTTGCT-----ACCTCTTACCCGCTCCCGGTTACTACTACTCTTATCCATCCC 1651
Oy      115  rOlEusErSerAlAheGlnSerProIlEProSerTyRGlNValSerProSerSers 135
Db      1652  CGGCGCACACCTCATACCCATCCCTGTGCGCACCTCTCTCTCTCTCTCTCTCTCTCTCTCA 1711
Oy      135  eRPhEProSerPro--SerArGlyGlupProAaAaAaAaAaAaAaAaAaAaAaAaAaAa 154
Db      1712  CTTACCATCCCTCTGTGACAGAGGCTCCCTCCCTCCGTCGTCGACACACATAC----- 1767
Oy      154  rOPheLeuArGAnGlyGlYlePro-----SerSerLeuProS 167
Db      1768  -----TCCTCTGTTCCTCCCTGCTTCCTCCCGGCGGACGATGAGTTCCTT 1813
Oy      167  eRLeuArGlyLeSerAnSerCybProValThrProProValSerSerProThSer---- 185
Db      1814  CTTAGCTGTACCACTCTCTTACAGGCGCTCCACAGGCGCTTTCGACATGACAGCAACT 1873
Oy      185  ----- 185
Db      1874  TTTCTCCAGCAATTGAATTGCTTAAAGGAAAGGAAAGGAAAGGAAAGGAGAGA 1933
Oy      186  --LyAsnProLybProLeuProAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 205
Db      1934  AAAAGAAACAAAGACATTAAAGACAGAG-----GAGGAGATGGCCATAG 1981
Oy      205  lAlYglnSerMetAlaSerPheAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 225
Db      1982  GAGAGAGGTTCTCTTAGTCAGATGAGGTTCT-----TCAGAGCAAGTCTTC 2032
Oy      225  rOThr 226
Db      2033  CCTCT 2037

```

RESULT 9

```

US-09-300-958A-32
; Sequence 32, Application US/09300958A
; Patent No. 6495319
; GENERAL INFORMATION:

```

```

APPLICANT: McCelland, Michael
APPLICANT: Welsh, John
TITLE OF INVENTION: Reduced Complexity Nucleic Acid Targets and Methods of
FILE REFERENCE: P-PH 3457
CURRENT APPLICATION NUMBER: US/09/300,958A
PRIOR APPLICATION NUMBER: 60/083,331
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/098,070
PRIOR FILING DATE: 1998-08-27
PRIOR APPLICATION NUMBER: 60/118,624
PRIOR FILING DATE: 1999-02-04
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 32
LENGTH: 3132
TYPE: DNA
ORGANISM: Homo sapiens
US-09-300-958A-32

Alignment Score:
Pred. No.: 0.0171 Length: 3132
Score: 137.50 Matches: 63.
Percent Similarity: 41.06% Conservative: 45
Beat Local Similarity: 23.95% Mismatches: 92
Query Match: 7.79% Indels: 63
DB: Gaps: 9

US-09-995-938A-7 (1-336) x US-09-300-958A-32 (1-3132)
Qy 2 Thiseraagcylalathserthser-----AlaAlaAlaAla 15
Db 1328 ACGAGCTCACCGCGACATCTCGCATCCACACAGCGCAGAACGCTTCAGTCCGCGCATCT 1387
Qy 16 AlaAlaAlaAlaAlaAla-ArgaGlyVaProSerTerTArgGluArgGluAsnArga 35
Db 1388 GCATGCGCAATTCACGCGCAGCGACCACTCACACCCACATCCGACCCACACAGCGG 1447
Qy 35 GARGGluArgaArgaArgaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 55
Db 1448 AAAAGCCCTTGCGCGCACATCTGTGAAGAAAGTTGGCCAGAGCGATGAACGACAGA 1507
Qy 55 YAspTyraAsnLeuProLyAHAsCyAAspAsnAsnGluValLeuAlaLeuCyValG 75
Db 1508 GGCATACCAAGATCCATCTCGCGCAGAAAGACAAAGACAGACAAAGCTGTGT----- 1562
Qy 75 wAlaGlyTrpValValGluGluAspGly-ThrThrTyraGlyAsGlyCyAlaYsProLeu 95
Db 1563 -----GGCCTCTTGCGGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1597
Qy 95 roGlyGluLeAlaGlyThrseerseraGValThrProTySerserGlnAsnGlnserP 115
Db 1598 CGTCCCGCGGTGCT-----ACCTCTTACCCGTCGCCGATTAACAACCTTATCCATCC 1651
Qy 115 roLeuSerSerAlaAheGlnSerProIleProSerTerTyGlnValserProSerSers 135
Db 1652 CGGCCACCACTCATACCATTCCTCTGTGCCCACTCTCTCTCTCTCTCTCTCTCTCTCTGA 1711
Qy 135 erPheProSerPro-----SerArgGlyGluProAsnAsnAsnMetSerSerThrPheP 154
Db 1712 CCTACCATCCCTGTGCAAGTGGCTTCCCTCCCGTCGGTGGCCACACAGTAC----- 1767
Qy 154 roPheLeuArgAsnGlyLysLePro-----SerSerLeuProS 167
Db 1768 -----TCTCTGTCTCCCTCCGCTTCTCCCGCCAGGTGACGACGCTTCTCTCT 1813
Qy 167 erLeuArgGILSerAsnSerCySproValThrProValSerSerProThrser----- 185
Db 1814 CCTCAGCTGTACCAACTCTTTCAGCGGCTTCCACAGGGCTTTTCGGACATGACAGCAACT 1873
Qy 185 ----- 185

```

```

Db      1874 TTTCCTCCAGGACCAATTGAATTGCTAAAGGGAGAAAGCGGAAAGGAAAAAGGAGCA 1933
Qy      186  --LysaenProlyseProleupProantPrngIuseRilleaLysGlnSerMetalailea 205
Db      1934 AAAAGAAACACACAAAGACTTAAAGAGACGAGG-----GAGCAGATGGCCATNG 1981
Qy      205  laLysGlnSerMetalaSerpheanantYrProphetyrAlaValSerAlaProAlaSerP 225
Db      1982 GAGGAGGAGGGTTCTCTTAAGTCAGATGAGAGTTT-----TCAGAGCCAAAGTCTCTC 2033
Qy      225  rothr 226
Db      2033 CCTCT 2037

RESULT 10
US-09-702-705-320
/ Sequence 320, Application US/09702705
/ Patent No. 6504010
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darlick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121,478C14
/ CURRENT APPLICATION NUMBER: US/09/702,705
/ CURRENT FILING DATE: 2000-10-30
/ NUMBER OF SEQ ID NOS: 1833
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 320
/ LENGTH: 3132
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-702-705-320

Alignment Scores:
Pred. No.: 0,0171 Length: 3132
Score: 137,50 Matches: 63
Percent Similarity: 41,06% Conservative: 45
Best Local Similarity: 23,95% Mismatch: 92
Query Match: 7,79% Indels: 63
DB: 4 Gaps: 9
US-09-995-938A-7 (1-336) x US-09-702-705-320 (1-3132)
Qy      2  ThiseraspGlyAlaThrSerThrSer-----AlaAlaAlaAlaAla 15
Db      1328 ACGAGCTCACCCGCGACATCCGCATCCACACAGCGCAGAAAGCCCTTCCAGTCCGCGCATCT 1387
Qy      16  AlAlaAlaAlaAlaAlaAla-ArgAaglylsProSerTrpArgGluArgGluAaenAaArg 35
Db      1388 GCATGCGCACTTACGCGCGGCGGACCACTTCCACCACCACCATCCGACCCACACAGGCG 1447
Qy      35  gArgGluArgArgArgArgAlaValAlaAlaAlaLysIleTyThrGlyLeuArgAlaGlnG 55
Db      1448 AAAAGCCCTTCCGCTCGGACATCTCTGTGAAGAAAGTTTGGCCAGAGCATGAAGCAAGA 1507
Qy      55  yAspTrpAsnLeuProLysHisCyAspAaenAaenGluValLeuLysAlaLeuCyValG 75
Db      1508 GGCATACCAAGATCCACTTCCGCGGACGAAGCAAGAAAGCAGACAAAGTGTGT----- 1567
Qy      75  uAlaGlyTrpValValGluGluAaenGly-ThrThrTyArgLysGlyCyValyProLeup 95
Db      1563 -----GGCTTTGGGCGCACCTCTCTCTCTTCTTCTTAC 1597
Qy      95  roGlyGluIleAlaGlyThrSerSerArgValThrProTySerSerGlnAaenGlnSerP 115

```

DB 1598 CGTCCCGGTTCT-----ACCTCTTACCGGCTCCCGGTTACTACTTATCATCC 1651
QY 115 rOlEuseRSeRAlaPheGInSeRProIleProSeRtyGInValSeRProSeRSeS 135
DB 1652 CGGACCACTCCATACCATCCCTGTCGCCACCTCTTCTCTCCGCTCCTCGA 1711
QY 135 eRPhEProSeRPro---SeRArgIyGluProAsnAsnAsnMetSeRSeRThrPhePheP 154
DB 1712 CTTACCCATCCCTGTGCACAGTGGCTTCCCTCCCGGTGGGACACGATAC--- 1767
QY 154 rOPheLeuArgAnGlyGlyIlePro-----SeRSeRLeuProS 167
DB 1768 -----TCCTCTGTTCCTCCCTTCCTCCCGGACGAGGACGACCTTCCCTT 1813
QY 167 eRLeuArgIleSeRAnSeRSeRtySProValThrProProValSeRSeRProThSeR--- 185
DB 1814 CTTACGCTGTCCCAACTCTCTTACAGCGCTCCACAGGCTTTCGACATGACAGCACT 1873
QY 185 ----- 185
DB 1874 TTTCTCCAGACCAATTGAAATTGCTTAAAGGAAAGGAAAGGAAAGGAGAGA 1933
QY 186 --LysAsnProLysProLeuProAsnTrpGluSerIleAlaLysGlnSeRMetaIleA 205
DB 1934 AAAAGAAACACAGAGACTTAAAGACAGAGAG-----GAGGAGATGGCCATAG 1981
QY 205 lAlYsGlnSeRMeCaLaseRpheAnTyRProPheTyRAlaValSeRAlaProAlaSeR 225
DB 1982 GAGAGAGAGGTTCTCTTAGTGCAGATGAGAGTTCTC-----TCAGAGCAAGTCTCTC 2032
QY 225 roThr 226
DB 2033 CCTCT 2037

RESULT 11

US-09-736-457-320
Sequence 320, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongrong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodea, Michael A.
APPLICANT: Panger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Derrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 320
LENGTH: 3132
TYPE: DNA
ORGANISM: Homo sapiens
US-09-736-457-320

Alignment Scores:

Pred. No.: 0.0171 Length: 3132
Score: 137.50 Matches: 63
Percent Similarity: 41.06% Conservative: 45
Best Local Similarity: 23.95% Mismatches: 92
Query Match: 7.79% Indels: 63
DB: 4 Gaps: 9

US-09-995-938a-7 (1-336) x US-09-736-457-320 (1-3132).

QY 2 ThrSeRAspGlyAlaThrSeRThSeR-----AlaAlaAlaAlaAla 15
DB 1328 ACAGCTCAACCGCCACATCCGATCCACACAGGCCAGAGCCCTTCCAGTCCGATCT 1387
QY 16 AlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 35
DB 1388 GCATCGCAACTTACCGCCGACGACGACGACGACGACGACGACGACGACGACGACGACG 1447
QY 35 gARgIuARgARgARgARgARgARgARgARgARgARgARgARgARgARgARgARgARgAR 55
DB 1448 AAAAGCCCTTGGCTCGACATCTGTGGAAGAAATTGCCAGAGCGATGAACGACAGA 1507
QY 55 yASpTyRAnLeuProLysHisCysAspAsnAsnAlaValLeuLysAlaLeuCysAlaG 75
DB 1508 GGCATACCAAGATCCACTTCCGCGAGAAAGACAAAGACAAAGACAAAGACTGTGT--- 1562
QY 75 uAlaGlyTrpValValGluGluAspGly-ThrThrTyRArgLysGlyCysLysProLeuP 95
DB 1563 -----GGCTCTTCGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1597
QY 95 rOGlyGluIleAlaGlyThrSeRSeRArgValThrProTyRSeRSeRglnAsnGlnSeR 115
DB 1598 CGTCCCGGTTGCT-----ACCTCTTACCGGCTCCCGGTTACTACTCTTATCCATCC 1651
QY 115 rOlEuseRSeRAlaPheGInSeRProIleProSeRtyGInValSeRProSeRSeS 135
DB 1652 CGGACCACTCCATACCATCCCTGTGTCGCCACCTCTCTCTCTCTCTCTCTCTCTCTCT 1711
QY 135 eRPhEProSeRPro---SeRArgIyGluProAsnAsnAsnMetSeRSeRThrPhePheP 154
DB 1712 CTTACCCATCCCTGTGCACAGTGGCTTCCCTCCCGGTGGGACACGATAC--- 1767
QY 154 rOPheLeuArgAnGlyGlyIlePro-----SeRSeRLeuProS 167
DB 1768 -----TCCTCTGTTCCTCCCTTCCTCCCGGACGAGTGCAGACTTCCCTT 1813
QY 167 eRLeuArgIleSeRAnSeRSeRtySProValThrProProValSeRSeRProThSeR--- 185
DB 1814 CTTACGCTGTCCCAACTCTCTTACAGCGCTCCACAGGCTTTCGACATGACAGCACT 1873
QY 185 ----- 185
DB 1874 TTTCTCCAGACCAATTGAAATTGCTTAAAGGAAAGGAAAGGAAAGGAGAGA 1933
QY 186 --LysAsnProLysProLeuProAsnTrpGluSerIleAlaLysGlnSeRMetaIleA 205
DB 1934 AAAAGAAACACAGAGACTTAAAGACAGAGAG-----GAGGAGATGGCCATAG 1981
QY 205 lAlYsGlnSeRMeCaLaseRpheAnTyRProPheTyRAlaValSeRAlaProAlaSeR 225
DB 1982 GAGAGAGAGGTTCTCTTAGTGCAGATGAGAGTTCTC-----TCAGAGCAAGTCTCTC 2032
QY 225 roThr 226
DB 2033 CCTCT 2037

RESULT 12

US-09-614-034-192
Sequence 192, Application US/09614034
Patent No. 6489307
GENERAL INFORMATION:
APPLICANT: PHILLIPS, M. IAN
APPLICANT: ZHANG, YUAN
TITLE OF INVENTION: ANTISENSE COMPOSITIONS TARGETED TO BETAL-ADRENOCEPTOR-SPECIFIC
FILE REFERENCE: 4300.013900
CURRENT APPLICATION NUMBER: US/09/614,034
CURRENT FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 09/152,717
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US99/21007
PRIOR FILING DATE: 1999-09-14
NUMBER OF SEQ ID NOS: 204

[illegible]

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 413-3000
: TELEFAX: (703) 413-2220
: TELE: 248855 OPAT UR
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1209 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1206
: US-08-314-309A-5

Alignment Scores:
Pred. No.: 0.00701 Length: 1209
Score: 134.50 Matches: 87
Best Similarity: 33.15% Conservative: 33
Best Local Similarity: 24.03% Mismatches: 122
Query Match: 7.62% Indels: 121
DE: 1 Gaps: 14

US-09-995-938A-7 (1-336) x US-08-314-309A-5 (1-1209)
OY 2 ThrSerApgly-...AlaThrSerThrSerAlaAlaAlaAlaAlaAlaAlaAla 20
Db 140 ACATTCCTGGTCTGCGACGACATCCACACGCGACCTCGACCGCCGCGGTA 199
OY 21 AlaArgArglyProSer----- 26
Db 200 TCGAGAAAGAGTCCACATCGACACCTTAAACGCTATGTGGCGGAGTTGACAGACTA 255
OY 27 ---TTPArgGluArgGluAsnAsnArgArgArgGluArgArgArgAlaAla-Al 45
Db 260 CCATTGAGAGAGATTCAGAACCAACCCGAGTGTCTCTGAGGTAGAGGAGGACACGACTCGC 319
OY 45 alysllelyrThrGlyLeuArgAlaGlnGlyAspTyrAsnLeuProLysHisCysAspAs 65
Db 320 ACCTCTTGACGAGACGAGGACGAGGAGAAATTACGACACGCCGCCCTCTCACCAGAACG 379
OY 65 nasngluValleuylalaLeuCysValGluAlaGlyTTPValValGluGluAspGlyTh 85
Db 380 GCGCCTGGGGCCTGGGACACATCTCTCATCGCAGCCTGCTCGACACAGCT-ACATCTAC 438
OY 85 rThrTyrArglyArgGlyCysLysProLeuProGlyGluIleAlaGlyThrSerSerArgVa 105
Db 439 GACGACTCGGCGCGGACGCGGACCTTACGCTTGTGTGACACGCGGCATCTCTGAGACT 498
OY 105 lThrProTyrSerSerGlnAsnGlnSerProLeuSerSerSerAlaPheGlnSerProIlePr 125
Db 499 CACA-----ACGAGTTCTCCGGCCG-G-CC 522
OY 125 oSerTyrGlnValSerProSerSerSerSerPheProSerProSerArgGlyGluProAs 145
Db 523 ATCAGGCGCTACACGCGCTCGCGCGGAGCAACGCGGACACCAACGCGCACCCAC 582
OY 145 nasn-----AsnMetSerSerThrPhePhePro-----Ph 155
Db 583 GTGGCTGGACCATTTGGCGGACGACCTAACGGCGTTGCCAAGAACCAACCATCATGCT 642
OY 155 eleuArgAsnGlyGly-----lleProSerSerLeuProSer----- 167
Db 643 GTCAAGGCTCTCGGGGAGATTTGAGCTCTACTTCCATATCTTTACGCGCTTCAACTGG 702
OY 168 -----LauArgIleSerAsnSerCysArgPro-- 175
Db 703 GCGGTGACGATATCATCAACAGGGGCGCGCAGAACACAGCGGCCATCAGCATGCTCGTG 762
OY 176 -----ValThrProValSe 181
Db 763 GGATGGCTACTCTTCTGCTTCAACAACGCGGTACACAGTCCACTACCGCGCGCTC 822

```


Db 976 -----C|A|A|C|C|C|A|T|C|A|C|C|G|G|C|T|C|C|A|T|C|C|A|A|C|T|-----G|G|G|C|C|G|C 1019
Qy 231 eH|s|T|h|r|L|e|u|A|A|T|h|r|L|e|P|r|o|G|I|u|C|y|A|s|p|G|l|u|S|e|r|A|s|p|S|e|r|T|h|r|V|a|l|A|s|p|S|e|r|G|I 251
Db 1020 T|C|C|A|G|C|T|T|C|A|G|C|A|C|T|A|G|G|G|C|T|C|G|T|G|T|C|G|A|C|A|T|C|T|----- 1056
Qy 251 yH|s|T|P|I|L|e|S|e|r|P|h|e|G|I|n|L|y|S|P|h|e|A|I|a|G|I|n|G|I|n|P|r|o|P|h|e|S|e|r|A|L|a|S|e|r|M|e|T|V|a|L|P|r 271
Db 1057 -----T|G|C|C|C|C|C|G|A|A|C|C|A|G|C|A|T|C|T|C|C|G|C|T|G|A|T|C|G|G|G|G|C|A|A|C|T|C|G|G|C 1106
Qy 271 oT|h|r|S|e|r|P|r|o|T|h|r|P|h|e|A|n|L|e|u|V|a|l|y|S|P|r|o|A|L|P|r|o|G|I|n|G|I|n|M|e|T|S|e|r|P|r|o|A|n|T|h|r|A|L 291
Db 1107 A|-----C|A|A|C|A|C|A|T|C|C|G|G|A|C|C|T|C|C|A|T|G|C|C|A|C|C|C|C|A|T|G|T|C|A|C|C|G|G|G|T|G|T|C 1160
Qy 291 aA|A 292
Db 1161 C|T|C|T 1164

Search completed: December 20, 2003, 05:51:30
Job time : 93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 20:59:46 ; Search time 800 Seconds
(without alignments)
1192.038 Million cell updates/sec

Title: US-09-995-938A-2
Perfect score: 2687
Sequence: 1 actttagtttgccttaattc.....atctttagtagataaaaaa 2687

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2211978 seqs, 1666101734 residues 4423956
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|---------|-------|---------------------|
| 1 | 2687 | 100.0 | 2687 | 13 | US-09-995-938A-2 |
| 2 | 2685.4 | 99.9 | 2687 | 13 | US-09-995-938A-1 |
| 3 | 1102 | 41.0 | 1456 | 13 | US-09-995-938A-5 |
| 4 | 1100.4 | 41.0 | 1456 | 13 | US-09-995-938A-4 |
| 5 | 713 | 26.5 | 3000 | 13 | US-09-995-938A-9 |
| 6 | 711.4 | 26.5 | 3000 | 13 | US-09-995-938A-3 |
| 7 | 535.6 | 19.9 | 1213 | 13 | US-09-995-938A-12 |
| 8 | 534 | 19.9 | 1213 | 13 | US-09-995-938A-11 |
| 9 | 149 | 5.5 | 831 | 10 | US-09-938-842A-385 |
| 10 | 121.4 | 4.5 | 413 | 10 | US-09-878-574-5002 |
| 11 | 109.6 | 4.1 | 284 | 10 | US-09-878-574-14631 |
| 12 | 71.4 | 2.7 | 3673778 | 13 | US-10-312-841-1 |
| 13 | 70 | 2.6 | 6145 | 13 | US-10-311-455-945 |
| 14 | 63.2 | 2.4 | 3673778 | 13 | US-10-312-841-2 |
| 15 | 62.6 | 2.3 | 9642 | 13 | US-10-311-455-330 |

Applications of other

| | | | | | | | |
|---|----|------|-----|---------|----|--------------------|--------------------|
| c | 16 | 61 | 2.3 | 3673778 | 13 | US-10-312-841-1 | Sequence 1, Appl1 |
| | 17 | 60.8 | 2.3 | 16093 | 13 | US-10-311-455-1377 | Sequence 1377, Ap |
| | 18 | 60.6 | 2.3 | 5608 | 13 | US-10-311-455-1111 | Sequence 1111, Ap |
| | 19 | 60.4 | 2.2 | 6070 | 13 | US-10-311-455-1652 | Sequence 1652, Ap |
| | 20 | 60.4 | 2.2 | 6070 | 13 | US-10-240-485-132 | Sequence 132, App |
| | 21 | 60.4 | 2.2 | 6954 | 13 | US-10-311-455-1364 | Sequence 1364, Ap |
| | 22 | 60.4 | 2.2 | 14006 | 13 | US-10-311-455-1931 | Sequence 1931, Ap |
| | 23 | 60.2 | 2.2 | 6283 | 13 | US-10-311-455-61 | Sequence 61, Appl1 |
| | 24 | 60.2 | 2.2 | 113515 | 13 | US-10-311-455-2147 | Sequence 2147, Ap |
| | 25 | 60 | 2.2 | 15373 | 13 | US-10-311-455-439 | Sequence 439, App |
| | 26 | 59.6 | 2.2 | 5649 | 13 | US-10-311-455-822 | Sequence 822, App |
| | 27 | 59.6 | 2.2 | 16891 | 13 | US-10-311-455-626 | Sequence 626, App |
| | 28 | 59.6 | 2.2 | 16891 | 13 | US-10-240-485-58 | Sequence 58, Appl1 |
| | 29 | 59.4 | 2.2 | 529 | 10 | US-09-983-965-2109 | Sequence 2109, Ap |
| | 30 | 59.4 | 2.2 | 20486 | 13 | US-10-240-485-164 | Sequence 164, App |
| | 31 | 59.2 | 2.2 | 6132 | 13 | US-10-311-455-462 | Sequence 462, App |
| | 32 | 59 | 2.2 | 446 | 10 | US-09-960-352-3400 | Sequence 3400, Ap |
| | 33 | 59 | 2.2 | 5195 | 13 | US-10-311-455-894 | Sequence 894, App |
| | 34 | 59 | 2.2 | 6351 | 13 | US-10-311-455-1720 | Sequence 1720, Ap |
| | 35 | 59 | 2.2 | 6944 | 15 | US-10-172-086-112 | Sequence 112, App |
| | 36 | 59 | 2.2 | 9539 | 13 | US-10-240-453-54 | Sequence 54, Appl1 |
| | 37 | 59 | 2.2 | 9539 | 15 | US-10-239-676-52 | Sequence 52, Appl1 |
| | 38 | 58.4 | 2.2 | 6061 | 13 | US-10-311-455-113 | Sequence 113, App |
| | 39 | 58.4 | 2.2 | 6189 | 13 | US-10-240-485-146 | Sequence 146, App |
| | 40 | 58.4 | 2.2 | 6668 | 13 | US-10-311-455-1669 | Sequence 1669, Ap |
| | 41 | 58.4 | 2.2 | 6668 | 13 | US-10-311-455-1670 | Sequence 1670, Ap |
| | 42 | 58.4 | 2.2 | 17721 | 13 | US-10-311-455-1701 | Sequence 1701, Ap |
| | 43 | 58.2 | 2.2 | 7490 | 13 | US-10-311-455-255 | Sequence 255, App |
| | 44 | 58.2 | 2.2 | 11944 | 13 | US-10-311-455-2159 | Sequence 2159, App |
| | 45 | 58 | 2.2 | 12705 | 13 | US-10-311-455-121 | Sequence 121, App |

ALIGNMENTS

RESULT 1
US-09-995-938A-2
Sequence 2, Application [US/09995938A
Publication No. US2003015002641
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
APPLICANT: ZHIYONG WANG
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
FILE REFERENCE: SALKINS.046A
CURRENT FILING DATE: US/09/995, 938A
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 2
LENGTH: 2687
TYPE: DNA
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-2

| | | | | |
|-----------------------|--------------|---|------------|-------------|
| Query Match | 100.0% | Score 2687 | DB 13 | Length 2687 |
| Best Local Similarity | 100.0% | Pred. No. 0 | | |
| Matches 2687 | Conservative | 0 | Mismatches | 0 |
| | | 0 | Indels | 0 |
| | | 0 | Gaps | 0 |
| QY | 1 | ACTTAGTTTGGCTTAATTCGAACTCTGATTCATTCGAAATGTTCCCACTGC | 60 | |
| DB | 1 | ACTTAGTTTGGCTTAATTCGAACTCTGATTCATTCGAAATGTTCCCACTGC | 60 | |
| QY | 61 | GTTGATGTTGGGTTCTGCTTTTAAAGACATTAAGTTAGTCTCTGCACCTTTCTAC | 120 | |
| DB | 61 | GTTGATGTTGGGTTCTGCTTTTAAAGACATTAAGTTAGTCTCTGCACCTTTCTAC | 120 | |
| QY | 121 | AAATTAACGCTATATCCAAATATTAATGATCATCATATATATGCGCGGACAC | 180 | |
| DB | 121 | AAATTAACGCTATATCCAAATATTAATGATCATCATATATATGCGCGGACAC | 180 | |
| QY | 181 | TGTTATGGAACCACTGTAACCCCTTTTCTTTTATGTTCAATGAATATACAGT | 240 | |
| DB | 181 | TGTTATGGAACCACTGTAACCCCTTTTCTTTTATGTTCAATGAATATACAGT | 240 | |

Db 181 TTGTTATGGGACAACTCGTAAACCCCTTTTCCCTTTATGTTCAATGACATATCAAGT 240
Qy 241 TTTGGTTATGAATACATAAATATATGATGAGACCCAGCAATTAATCCAAAATTTGGATATTA 300
Db 241 TTTGGTTATGAATACATAAATATATGATGAGACCCAGCAATTAATCCAAAATTTGGATATTA 300
Qy 301 GATATCAAGGCTTAAATCAACATGTAACCAACTTAATACCTTATAGAACTATGAAT 360
Db 301 GATATCAAGGCTTAAATCAACATGTAACCAACTTAATACCTTATAGAACTATGAAT 360
Qy 361 GGTATTCACCAATCTTTATATCATTTGTAGGTACGAAAGAGTAAAGAGAGAGAGAGAG 420
Db 361 GGTATTCACCAATCTTTATATCATTTGTAGGTACGAAAGAGTAAAGAGAGAGAGAGAG 420
Qy 421 CAGGTATACATACATTAATCAGAGCAAAAGTAGTCAAAATTTGTTCTTAAGTAGATTTG 480
Db 421 CAGGTATACATACATTAATCAGAGCAAAAGTAGTCAAAATTTGTTCTTAAGTAGATTTG 480
Qy 481 TATGCAAAAAAAGAGTATATTTTATGCAATATATCATATATGATTTATCAACGG 540
Db 481 TATGCAAAAAAAGAGTATATTTTATGCAATATATCATATATGATTTATCAACGG 540
Qy 541 CTTCACCTATTAATTAATTTTAAACCAATGAATGCTGCTTTTGGGTACACA 600
Db 541 CTTCACCTATTAATTAATTTTAAACCAATGAATGCTGCTTTTGGGTACACA 600
Qy 601 AGTAAAGGACCCACACTGAGAAAGCGTCCACGTGCTCCTCCTTTCTTTCTCTGTA 660
Db 601 AGTAAAGGACCCACACTGAGAAAGCGTCCACGTGCTCCTCCTTTCTTTCTCTGTA 660
Qy 661 TTATTTGGTCATTTACTCATTTTACTACTACAGAAAAAAGATTAGAACATTAAC 720
Db 661 TTATTTGGTCATTTACTCATTTTACTACTACAGAAAAAAGATTAGAACATTAAC 720
Qy 721 ACACGTTACTAAGCGTAGTATCTCTGACCTTAACATACACCTTATATTCACCTCA 780
Db 721 ACACGTTACTAAGCGTAGTATCTCTGACCTTAACATACACCTTATATTCACCTCA 780
Qy 781 CGTAACTCAACCCCTTCAAAACCATGTAATTTACAGGTGAGCGATGATACACAGAACNA 840
Db 781 CGTAACTCAACCCCTTCAAAACCATGTAATTTACAGGTGAGCGATGATACACAGAACNA 840
Qy 841 TGATTTCTTAATATGAATCTCATATGTAACACACACACGCCAATTTTATCACTTGA 900
Db 841 TGATTTCTTAATATGAATCTCATATGTAACACACACACGCCAATTTTATCACTTGA 900
Qy 901 TGAATAAATAATTAATTTGTTGAGAAAGAAAGAGATTTCTTCTTCCGATCCAGC 960
Db 901 TGAATAAATAATTAATTTGTTGAGAAAGAAAGAGATTTCTTCTTCCGATCCAGC 960
Qy 961 GAAAGAAAAGCGTATTTCTCGTAGACATTAATTTCTCACTCCTCTCTTCTTCTTCAAT 1020
Db 961 GAAAGAAAAGCGTATTTCTCGTAGACATTAATTTCTCACTCCTCTCTTCTTCTTCAAT 1020
Qy 1021 CAGTCTACGTTTCAACAATCTTTCAACCCACTAATCAAAAGCTCTCTCCGGAGTTTCAG 1080
Db 1021 CAGTCTACGTTTCAACAATCTTTCAACCCACTAATCAAAAGCTCTCTCCGGAGTTTCAG 1080
Qy 1081 GGGTTGTGTTGGTTTTCGCCATGATGCTTCCGATGAGAGTACGACATCAGCAGCTGC 1140
Db 1081 GGGTTGTGTTGGTTTTCGCCATGATGCTTCCGATGAGAGTACGACATCAGCAGCTGC 1140
Qy 1141 AGCTCGCGCGCGGAGCAGCGCGAGAGAGAGAGCGTGTGAGAGAAAGGAGAAATTA 1200
Db 1141 AGCTCGCGCGCGGAGCAGCGCGAGAGAGAGAGCGTGTGAGAGAAAGGAGAAATTA 1200
Qy 1201 TCGAG 1260
Db 1201 TCGAG 1260
Qy 1261 TCAAGGTATTAATTTTGGCTTAACATTTGATTAATTAAGAGTCTTAAGCTCTTTG 1320
Db 1261 TCAAGGTATTAATTTTGGCTTAACATTTGATTAATTAAGAGTCTTAAGAGTCTTTG 1320

Qy 1321 TGTGAAAGCTGTGGGTGTTGTAAGAGATGTAATCTATATGCAAGGTGAAGACTTT 1380
Db 1321 TGTGAAAGCTGTGGGTGTTGTAAGAGATGTAATCTATATGCAAGGTGAAGACTTT 1380
Qy 1381 CTCGATTTTTCGAGCTGAGCTGTTGTTTATGATGTTTGTGAGTTGAACTCGAAT 1440
Db 1381 CTCGATTTTTCGAGCTGAGCTGTTGTTTATGATGTTTGTGAGTTGAACTCGAAT 1440
Qy 1441 CGTTGATTTCAATGTGGTTAAATGGGTTGAAATTTGAGAAATTTGAGGTTTCTCAAAG 1500
Db 1441 CGTTGATTTCAATGTGGTTAAATGGGTTGAAATTTGAGAAATTTGAGGTTTCTCAAAG 1500
Qy 1501 TGAATTTGAATCATCAGAAACATATGATGATCTTAATTTCAAGTGAATTTAGGTT 1560
Db 1501 TGAATTTGAATCATCAGAAACATATGATGATCTTAATTTCAAGTGAATTTAGGTT 1560
Qy 1561 TTTCTTTCTAATTTTGAATATTAATGATGCTTAAGTCTTAATTTATGATGAT 1620
Db 1561 TTTCTTTCTAATTTTGAATATTAATGATGCTTAAGTCTTAATTTATGATGAT 1620
Qy 1621 CTGAGTCCAAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 CTGAGTCCAAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Qy 1681 TTGGTTATTTGTTTGGCTTTTGGAGTATCAGGATGACAGGCTTTACCTGGTGAATA 1740
Db 1681 TTGGTTATTTGTTTGGCTTTTGGAGTATCAGGATGACAGGCTTTACCTGGTGAATA 1740
Qy 1741 GCTGGAGTCTATCTCGAGTAACTCCATATTCATATACAGAAACAGAGCCCTTTTCACTA 1800
Db 1741 GCTGGAGTCTATCTCGAGTAACTCCATATTCATATACAGAAACAGAGCCCTTTTCACTA 1800
Qy 1801 GCTTTCAAAGTCCCATCCCATCTTAACCAAGTTAACCGCTCTTCTTCAATTTCCGAGT 1860
Db 1801 GCTTTCAAAGTCCCATCCCATCTTAACCAAGTTAACCGCTCTTCTTCAATTTCCGAGT 1860
Qy 1861 CTTCTCGGGGTGAACCAATTAACCAATGCTCTTATCATTTCTTCTTCTTCTTCAAGAT 1920
Db 1861 CTTCTCGGGGTGAACCAATTAACCAATGCTCTTATCATTTCTTCTTCTTCTTCAAGAT 1920
Qy 1921 GGTGGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
Db 1921 GGTGGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
Qy 1981 CCGGTCTCATGCGGACTTCTTAAGAACCGAAACGTTGCTTAATCTGGAATCTATCGCT 2040
Db 1981 CCGGTCTCATGCGGACTTCTTAAGAACCGAAACGTTGCTTAATCTGGAATCTATCGCT 2040
Qy 2041 AAGCAATCATGCGGACTTCTTAAGAACCAATTAAGAGTCTTAAATTTATCTTATGCG 2100
Db 2041 AAGCAATCATGCGGACTTCTTAAGAACCAATTAAGAGTCTTAAATTTATCTTATGCG 2100
Qy 2101 GTTTCTGACCTGCTAGTCTGAGACATATGCGACCAAGTTTCAATCTTCTTCTTCTTCT 2160
Db 2101 GTTTCTGACCTGCTAGTCTGAGACATATGCGACCAAGTTTCAATCTTCTTCTTCTTCT 2160
Qy 2161 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db 2161 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Qy 2221 TTTGCAACAACAAGCCATTTCTGCTCTTATGAGTCCAACTCTTCAATCTTCAATCTT 2280
Db 2221 TTTGCAACAACAAGCCATTTCTGCTCTTATGAGTCCAACTCTTCAATCTTCAATCTT 2280
Qy 2281 GTGAAACCTGCGCTCAGCAGATGCTTCAAAATATGCTGCTTCCAAAGATTTGCTCA 2340
Db 2281 GTGAAACCTGCGCTCAGCAGATGCTTCAAAATATGCTGCTTCCAAAGATTTGCTCA 2340
Qy 2341 AGCTCTGATTTAATTTGAGATGAGCAAGTTAAACCTGGAGAGAGAGAGATACAT 2400
Db 2341 AGCTCTGATTTAATTTGAGATGAGCAAGTTAAACCTGGAGAGAGAGAGATACAT 2400

QY 2401 GATGGGATGAGGATCTTGAAGCTTGAAGAGGCGGTTGACAT 2460
DB 2401 GATGGGATGAGGATCTTGAAGCTTGAAGAGGCGGTTGACAT 2460
QY 2461 AAACAATGAGCAACCAATGAGCATGTCATTTGAATATGAGAACTATCTCTTGAG 2520
DB 2461 AAACAATGAGCAACCAATGAGCATGTCATTTGAATATGAGAACTATCTCTTGAG 2520
QY 2521 TATTTCTCTGTCGACAGTATTTGATGATCTTATGAACTCTATGTTCTTCACTTAT 2580
DB 2521 TATTTCTCTGTCGACAGTATTTGATGATCTTATGAACTCTATGTTCTTCACTTAT 2580
QY 2581 TATTCGAATATGTCGACCAAGCTTCTCCATGGAAGATGAGGTGATAGGTAT 2640
DB 2581 TATTCGAATATGTCGACCAAGCTTCTCCATGGAAGATGAGGTGATAGGTAT 2640
QY 2641 TCAATTCATGATTTGGTTTCAAAAGCATTTATTTGATATAAAAA 2687
DB 2641 TCAATTCATGATTTGGTTTCAAAAGCATTTATTTGATATAAAAA 2687

RESULT 2

US-09-995-938A-1
Sequence 1, Application US/09995938A
Publication No. US20030150026A1
GENERAL INFORMATION:
APPLICANT: JOANNE CHORY
APPLICANT: ZHIYONG WANG
TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
FILE REFERENCE: SALKINS.046A
CURRENT APPLICATION NUMBER: US/09/995.938A
CURRENT FILING DATE: 2001-11-27
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 1
LENGTH: 2687
TYPE: DNA
ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938A-1

Query Match 99.9%; Score 2685.4; DB 13; Length 2687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2686; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTTAGTTGGCCATTCATGAAACCTGATTCATTCGAATGTTCCCACTGCG 60
DB 1 ACTTAGTTGGCCATTCATGAAACCTGATTCATTCGAATGTTCCCACTGCG 60
QY 61 GTTGATGTTGGGTTCCCTGCTTTTAAAGACATTAGTACATGCTCTGCACTTTTAC 120
DB 61 GTTGATGTTGGGTTCCCTGCTTTTAAAGACATTAGTACATGCTCTGCACTTTTAC 120
QY 121 AAAATAAACGTCATATCCAAAAATTTACATGATCATCATATATATGCGCGGAAC 180
DB 121 AAAATAAACGTCATATCCAAAAATTTACATGATCATCATATATATGCGCGGAAC 180
QY 181 TTGTTATGAGCAAACTGCTAAACCTTTTCTTTTATGTTCAATGATATCAAGT 240
DB 181 TTGTTATGAGCAAACTGCTAAACCTTTTCTTTTATGTTCAATGATATCAAGT 240
QY 241 TTTGTTATGAT 300
DB 241 TTTGTTATGAT 300
QY 301 GATTAATTAAGTTAAATCAATGTAACCAATTAATATATATATATATATATATAT 360
DB 301 GATTAATTAAGTTAAATCAATGTAACCAATTAATATATATATATATATATATAT 360
QY 361 GGTATTCACCAATCTTATATATATATATATATATATATATATATATATATATAT 420
DB 361 GGTATTCACCAATCTTATATATATATATATATATATATATATATATATATATAT 420

QY 421 CAGTGTACATACAACTATATGAGCAAAAGTAGTCAAAATGTTCTTAAAGTGATTTG 480
DB 421 CAGTGTACATACAACTATATGAGCAAAAGTAGTCAAAATGTTCTTAAAGTGATTTG 480
QY 481 TATGCAAGAAAAGTATATATTTTAAATGATATATATATATATATATATATATATAT 540
DB 481 TATGCAAGAAAAGTATATATTTTAAATGATATATATATATATATATATATATATAT 540
QY 541 CTTCACGTAT 600
DB 541 CTTCACGTAT 600
QY 601 AGTAAGGAGCAACATGAGAAAGCGTCCCATGTCCTCTCTCTCTCTCTCTCTCTCT 660
DB 601 AGTAAGGAGCAACATGAGAAAGCGTCCCATGTCCTCTCTCTCTCTCTCTCTCTCT 660
QY 661 TTTATGTCATATCTCATTTTACATATCTCAGAAAAAAGATTTAGAACATTAAC 720
DB 661 TTTATGTCATATCTCATTTTACATATCTCAGAAAAAAGATTTAGAACATTAAC 720
QY 721 ACAGTTACTATAGGTATATATGCTCTGACCTTAACTACACCTTATATATCTCTCA 780
DB 721 ACAGTTACTATAGGTATATATGCTCTGACCTTAACTACACCTTATATATCTCTCA 780
QY 781 CGTATCTACCTCTCCAAAAACATGATTTTACATGAGAGATGATACACAAACA 840
DB 781 CGTATCTACCTCTCCAAAAACATGATTTTACATGAGAGATGATACACAAACA 840
QY 841 TGATTTCTTATATGATCAATGATGATGAAACACACACAGACCAATTTTATACATTA 900
DB 841 TGATTTCTTATATGATCAATGATGATGAAACACACACAGACCAATTTTATACATTA 900
QY 901 TGAATAAATATATATATATTTTGTGAGAAAGAAAGATTTCTTCTGATTCACAC 960
DB 901 TGAATAAATATATATATATTTTGTGAGAAAGAAAGATTTCTTCTGATTCACAC 960
QY 961 GAAGAAAAGGATTTCTCTGAGCACTTACTCTCTCTCTCTCTCTCTCTCTCTCTCAT 1020
DB 961 GAAGAAAAGGATTTCTCTGAGCACTTACTCTCTCTCTCTCTCTCTCTCTCTCTCAT 1020
QY 1021 CAGTGTACATACAACTATATGAGCAAAAGTAGTCAAAATGTTCTTAAAGTGATTTG 1080
DB 1021 CAGTGTACATACAACTATATGAGCAAAAGTAGTCAAAATGTTCTTAAAGTGATTTG 1080
QY 1081 GGGTTGTTGTTGTTTCCGATGATCTTGGATGAGCTGACATTCAGCAGCTGCG 1140
DB 1081 GGGTTGTTGTTGTTTCCGATGATCTTGGATGAGCTGACATTCAGCAGCTGCG 1140
QY 1141 AGCTGCGGCGGCGGAGAGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTA 1200
DB 1141 AGCTGCGGCGGCGGAGAGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTA 1200
QY 1201 TCGAG 1260
DB 1201 TCGAG 1260
QY 1261 TCAAGGTATATATATTTGCTTAAATGATATATATATATATATATATATATATATAT 1320
DB 1261 TCAAGGTATATATATTTGCTTAAATGATATATATATATATATATATATATATATAT 1320
QY 1321 TGTGTAAGCTGTTGGTTGTTGTAAGAGATGTTTACTTATATGCAAGGTGAAGCTTT 1380
DB 1321 TGTGTAAGCTGTTGGTTGTTGTAAGAGATGTTTACTTATATGCAAGGTGAAGCTTT 1380
QY 1381 CTTCATTTTTCACATCTGAGCTGTTTATATATATATATATATATATATATATATAT 1440
DB 1381 CTTCATTTTTCACATCTGAGCTGTTTATATATATATATATATATATATATATATAT 1440
QY 1441 CGTTGATTTCAATTTGTTTAAATGAGGTTGAATCTGAGATTTTGAAGGTTTCTCAAG 1500
DB 1441 CGTTGATTTCAATTTGTTTAAATGAGGTTGAATCTGAGATTTTGAAGGTTTCTCAAG 1500
QY 1501 TGAATTTGATATATAGAAACTATATGATGATCTATTTCTCAAACTGAATTAAGGTT 1560


```

QY 1611 ANGTATGATCTGTGTCGAAGTCATGATGTTGTTCTTGTACCTGATGATGAT 1670
DB 1845 ATTGA-----GATCTCTCTTGAAGTTGA 1874
QY 1671 TGATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1730
DB 1875 TTGTGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1934
QY 1731 TGTGAGATAGTGGGACCTTCATCTGAGTAATCTCATATTCATACAGAACGAGCCC 1790
DB 1935 TGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1994
QY 1791 TCTTTATCAGCCTTTCAAGTCCCATCCCATCTTACCAAGTACCCGCTTCTTATC 1850
DB 1995 TCTTTCTCAGCTTTTATAGCCCATCTTATCTTACCAAGTACCTTCTTCTTCTC 2054
QY 1851 ATTCGAGTCTCTGCGGGTGAACCAATACATGTCCTTACATTTCCCTT 1910
DB 2055 ATTCGAGTCTCTGCGGGTGAACCAATACATGTCCTTACATTTCCCTT 2111
QY 1911 CCTCAGAAATGTTGGAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1970
DB 2112 CCTCAGAAATGTTGGAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2171
QY 1971 AGTTACCCACCGGTCTCATGCGGACTTTTAAAGAACCCGAAACCGTTGCTTAC 2030
DB 2172 TGTCATCTCACCACTGTCATCCCAACTTAAAGAACCCGAAACCGTTGCTTAC 2231
QY 2031 ATCTATGGCTAAGATTCATAGGCACT--TGCTAAACATCAATAGGCTCTTTAATTA 2087
DB 2232 ATCTATGGCTAAGATTCATAGGCACT--TGCTAAACATCAATAGGCTCTTTAATTA 2291
QY 2088 TCCTTTCTATGCGGTTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2147
DB 2292 CCCGTTTATGCGGTTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2351
QY 2148 GGTCTATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2207
DB 2352 GGTCTATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2411
QY 2208 AAGCTTTCAGAGTTGTCAGAACCAAGCAATCTCTGCTGCTGCTGCTGCTGCTGCT 2267
DB 2412 AAGCTTTCAGAGTTGTCAGAACCAAGCAATCTCTGCTGCTGCTGCTGCTGCTGCT 2471
QY 2268 TACCTTCAATCTTGAAGACCTGCGCTCAGCAGATGCTCCAAATCTGCTGCTTCA 2327
DB 2472 TACCTTCAATCTTGAAGACCTGCGCTCAGCAGATGCTCCAAATCTGCTGCTTCA 2531
QY 2328 AGAGATTTGTCAGAACCTGAGTTTAAATTTGAGAAATGCAAGTTAAACCTGCGAAG 2387
DB 2532 AGAGATTTGTCAGAACCTGAGTTTAAATTTGAGAAATGCAAGTTAAACCTGCGAAG 2591
QY 2388 AGAGAGATTAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2447
DB 2592 GAGAGGATTCATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2651
QY 2448 TCGTGTGATGATTAACAATAGCAAAACCAATGCGATGTC 2490
DB 2652 TCATAGTTGATGATTAACAATAGCAAAACCAATGCGATGTC 2694

```

```

; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: ARABIDOPSIS THALIANA
US-09-995-938a-3

Query Match      26.5%; Score 711.4; DB 13; Length 3000;
Best Local Similarity 71.1%; Pred. No. 1.6e-161;
Matches 1097; Conservative 0; Mismatches 376; Indels 70; Gaps 9;

QY 958 AGCGAAGAAAGGATATCTCTGAGCACTAATCTCTCTCTCTCTCTCTCTCTCTCTCT 1017
DB 1212 AGAAGAAAGGATATCTCTGAGCACTAATCTCTCTCTCTCTCTCTCTCTCTCTCT 1267
QY 1018 CATGATCTACGTTACACAAATCTTCAACCCATTTCAAGCTCTCTCGGAAGTTTC 1077
DB 1268 CAGCTGAATCAAAATACCATTTCCATTTTAACTGTTGATGTTGTTGAGAGTTGA 1327
QY 1078 GAGGAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1137
DB 1328 AGGAAGAG-----ATGACCTTACGAGCAACGTCGAGCT--CAAGC 1368
QY 1138 TGCACTGCGGCGGCGAGCAGCGGAGGAGAGAGCCGTCGTGAGAGAGAGAGAGAG 1197
DB 1369 TGCACTGCGGCGGCGAGCAGCGGAGGAGAGAGCCGTCGTGAGAGAGAGAGAGAG 1428
QY 1198 TAAATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
DB 1429 CAATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1488
QY 1258 AGCTCAAGTATTAATTTGCTTAAACATTTGATTAATTAATGAGTCTTAAAGCTCT 1317
DB 1489 AGCTCAAGTATTAATTTGCTTAAACATTTGATTAATTAATGAGTCTTAAAGCTCT 1548
QY 1318 TTGTGTTGAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1377
DB 1549 TTGTGTTGAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1608
QY 1378 TTTCTCATTTTTTCAGATCTGAGCTTGTATTAATGTTTGAATGTTGAATCTGA 1437
DB 1609 CAAGCAGATTTTTC-----TTAGATCTGATTAATGTTGAATGTTGAATCTGA 1664
QY 1438 ATTCGTTGATTTCAATGTTGTTTAAATGAGTTGAAT--CTGAGATTTGAGGTTTC 1494
DB 1665 AGATTTGATTTGAATTTGTTCAAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1724
QY 1495 TC-----AAAGTGAATTTGAATCATCAGAAACATGATGATCTGATTTCTCAAAGTGA 1550
DB 1725 GCTGAGAGAGCTCAAGTTTATTTATTTAAATTTGATGAGATTTCTTGGGAGAAAGTAC 1784
QY 1551 TTTATGAGTTTCTCTTCAATTTAGTTTAAATTTGATTTGATTTGATTTGATTTGATTT 1610
DB 1785 TTTAGGTTCTTCTTCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1844
QY 1611 ATGATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1670
DB 1845 ATTGA-----GATCTCTCTTGAAGTTGA 1874
QY 1671 TGATGTTGATTTGTTATTTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1730
DB 1875 TTGTGACATTTGATTTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1934
QY 1731 TGTGAGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1994
DB 1935 TGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2054
QY 1791 TCTTTATCAGCCTTTCAAGTCCCATCCCATCTTACCAAGTACCCGCTTCTTCTTATC 1850
DB 1995 TCTTTCTCAGCTTTTATAGCCCATCTTATCTTACCAAGTACCCGCTTCTTCTTCTC 2054
QY 1851 ATTCGAGTCTCTGCGGGTGAACCAATACATGTCCTTACATTTCCCTT 1910

```

RESULT 6
US-09-995-938a-3

; Sequence 3, Application US/0995938A
; Publication No. US20030150026A1
; GENERAL INFORMATION:
; APPLICANT: JOANNE CHORY
; APPLICANT: ZHIYONG WANG
; TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID
; FILE REFERENCE: SALKINS.046A
; CURRENT APPLICATION NUMBER: US/09/995,938A
; CURRENT FILING DATE: 2001-11-27

Best Local Similarity 82.0%; Pred. No. 9.1e-119;
Matches 641; Conservative 0; Mismatches 135; Indels 6; Gaps 2;

| | | | |
|----|------|---|------|
| QY | 1712 | AGGAGTCAAGACCTTTACCTGGTGAATAGCTGGACCTTCATCTCGAGTAATCCATATT | 1771 |
| Db | 409 | AGGAGACACAAAGCCTCTACCTGGTACATGGCTGGATATCTTCTCGAGCAACTCTTACT | 468 |
| QY | 1772 | CATCAGCAACAGAGCCCTCTTTCATCAGCTTTCAAAGTCCATCTTACCAG | 1831 |
| Db | 469 | CTTCCCATTAACCAAGTCTCTTCTTCTTCCACTTTTGATAGCCCATCTTATCTTACCAG | 528 |
| QY | 1832 | TTAGCCGCTCTTCTTCAATATCCCGAGTCTTCTCGGGGAAACAAATPACAACTGT | 1891 |
| Db | 529 | TCATCTCTTCTCTTCTTCTTATCCCGAGTCTTCTCGAGTTGGATGCACACAATAT-- | 586 |
| QY | 1892 | CCCTACATCTTCCCTTCTCTCAGAAATGGTGGCATCTTCTTCTTCTTCCCTCA | 1951 |
| Db | 587 | -CTCCACATCTTCCCTTCTCTCAGAAATGGTGGATCTTCTTCTTCTCTCTCACTTA | 645 |
| QY | 1952 | GAATCTCAACAGTTTGTTCAGTTAACCCGCGGTCTCATGGCCGACTTCTPAAAGCCCA | 2011 |
| Db | 646 | GAATCTCAACAGTCTCTGTCTACTCCACAGTGTATCTCCCACTTCTTAAAGAACCCA | 705 |
| QY | 2012 | AACGGTGGCCTAACGGGGAATCTATGGTAGCAATCCATGGGCAT--TGCATAACAT | 2068 |
| Db | 706 | AACCATGGCTACTTGGGAATCTTTTACCAAAACATTCATTCATGGCTGCTTAAGT | 765 |
| QY | 2069 | CAATGGCGTCTTTTAATTATCTTTCTATAGCGGTTTCTGACCTGTCTATGTCGACATC | 2128 |
| Db | 766 | CAATGACTCTTTGAACTACCCGTTTATGCGGTGTCTGACCTGGCTCTACTCATC | 825 |
| QY | 2129 | GCCACCACTTTCATACCTGGCTACTATACCTGAAATGATGATGATGACTCTTCCACTG | 2188 |
| Db | 826 | ATCGCACTTCCATAGCTCCGGCTACTATACCTGAATGATGAAGTGTGACTCTTCCACTG | 885 |
| QY | 2189 | TTGATTTCTGTGATCTTGATTAAGCTTTGAGAAGTTTGACAACACAGCCATCTCTGCT | 2248 |
| Db | 886 | TTGATTTCTGTGATCTTGATTAAGCTTTGAAAAGTTTGACAAACACAGCCATCTCTGCT | 945 |
| QY | 2249 | CTATGTGCCAACCCTCTCTTACCTTCAATCTTGTAACCTGGCGCTTAGAGATGTCT | 2308 |
| Db | 946 | CTATGTGCCAACCCTGGCTACTCTTCAATCTGTGTAAACCTGGCACACAGCAATGTCTC | 1005 |
| QY | 2309 | CAAAATACGTCTGCCCTTCCAAAGATTTGATGCAAAAGCTCGAGTTTAAATTGGAATAGCC | 2368 |
| Db | 1006 | CAAAATACGTCTGCCCATCTCCAAAGATTTGTCAAAGCTCCGAGTTTAAATTGGAACACCC | 1065 |
| QY | 2369 | AAGTTAAACCCCTGGGAGAGAGAGAGATACATGATGTGGGATATGGAGATCTTGAGCTTA | 2428 |
| Db | 1066 | AAGTTAAAGCCATGGGAAGGAGAGAGATCCATGATGTGGCTATGGAGATCTTAAGAGCTCA | 1125 |
| QY | 2429 | CACTTGGAATGGGAAGGCTCTGTGTTGACATPAAACACTAGGCCAAACCAATGGCATG | 2488 |
| Db | 1126 | CGCTTGAAGATGTAAAGCTCATAGTTGATGAAGATATACATGAACCTGTATTATCATG | 1185 |
| QY | 2489 | TC 2490 | |
| Db | 1186 | TC 1187 | |

RESULT 9
 US-09-938-842A-385
 / Sequence 385, Application US/09938842A
 / Patent No. US20020160378A1
 / GENERAL INFORMATION:
 / APPLICANT: Harper, Jeff
 / APPLICANT: Kreps, Joel
 / APPLICANT: Wang, Xun
 / APPLICANT: Zhu, Tong
 / TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 / TITLE OF INVENTION: SAME, AND METHODS OF USE
 / FILE REFERENCE: SRIPI300-3
 / CURRENT APPLICATION NUMBER: US/09/938, 842A

```

: CURRENT FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 385
: LENGTH: 831
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-385

```

| | Query Match | Similarity | 5.5%; | Score 149; | DB 10; | Length 831; |
|----|-------------|--|--------------|--------------------|------------|-------------|
| | Best Local | Similarity | 76.8%; | Pred. No. 1,76-25; | | |
| | Matches | 182; | Conservative | 0; | Mismatches | 55; |
| | | | | | Indels | 0; |
| | | | | | Gaps | 0; |
| QY | 1134 | CAGCTGACGCTGCGCGCGCGGACGAGTGGCGAGATAGAGAAACCGCTCGTGTGAGAGAAAGG | | | | 1199 |
| DB | 5 | CGGATGATAGGAGGAGATCAACCGCGCGCGAGAGAGATGCGACGTGGAGAGAAAG | | | | 64 |
| QY | 1194 | AGAAATATCGGAGAGAGAAAGCGAGAGAGCTGTAGCTGTGCGAAGATATACACTGGGC | | | | 1253 |
| DB | 65 | AGAACACACAGAGAGAGAAAGAGAAAGAACGAAATAGCAGCTTAAATCTTCACCGGAC | | | | 124 |
| QY | 1254 | TTAGAGCTCAAGGTGATTAATTAATTGCTTAAACATTTGTGATTAATGAAAGTCTTAAAG | | | | 1311 |
| DB | 125 | TTAGATCTCAAGCAATTAATTAATCTCTTAAACATGTGACACAAATGMACTCTCTCAAG | | | | 184 |
| QY | 1314 | CTCTTTGTGTTGAAGCTGTGGGTGTGTGAGAAAGATGTGACTACTATGCAAG | | | | 1370 |
| DB | 185 | CTCTTTGTCTTGAAGCTGTGTGATCGTTCAATGAAATGSCACCACTTAATGAAAG | | | | 241 |

```

RESULT 10
US-09-878-574-5002
/ Sequence 5002, Application US/09878574
/ Patent No. US20020110548A1
/ GENERAL INFORMATION:
/ APPLICANT: Byrum, Joseph R.
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Thompson, Michael D.
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(15401)B
/ CURRENT APPLICATION NUMBER: US/09/878,574
/ CURRENT FILING DATE: 2001-12-21
/ PRIOR APPLICATION NUMBER: 09/333,535
/ PRIOR FILING DATE: 1999-06-14
/ NUMBER OF SEQ ID NOS: 15775
/ SEQ ID NO 5002
/ LENGTH: 413
/ TYPE: DNA
/ ORGANISM: Glycine max
/ OTHER INFORMATION: Clone ID: LIB3028-009-Q1-B1-D2
US-09-878-574-5002

```

| | Query Match | Similarity | 69.8% | Score 121.4 | DB 10 | Length 413 | |
|----|-------------|---|-------|-------------------|--------|------------|--------|
| | Best Local | | | Prod. No. 5.3e-19 | | | |
| | Matches 164 | Conservative | 0 | Mismatches 71 | Indels | 0 | Gaps 0 |
| Qy | 1136 | GCTCAGCTGCGCGCGCGCAGCAGCGCCGAGGAAAGGAG | | | | | 1195 |
| Db | 92 | GCCGGAACATGACCGCGCGCGATCCACGGGAGGTTGCCGAGTGGAGAGAGAG | | | | | 151 |
| Qy | 1196 | AATATCGAGAGAGAGAAAGCGAGAAAGCTTACTGCGAAGATTACAGGGCT | | | | | 1255 |
| Db | 152 | AACAAACAGGAGAGAGAGAAAGCGAAGCCGATTGACGTAAAGATTACATGGCTT | | | | | 211 |
| Qy | 1256 | AGAGCTCAAGGATTAATTAATTTGGCTAAACATGTGTATTAATTAAGTCTTAAAGCT | | | | | 1315 |
| Db | 212 | CGAGCCAGGAACTTCAAGCGCTCCAGAGCATCTGCACAAACAGAGTCTGTAAAGT | | | | | 271 |

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 20:38:22 ; Search time 9603 Seconds

(without alignments) 11446.857 Million cell updates/sec

Title: US-09-995-938a-2

Perfect score: 2687

Sequence: 1 attcattgttcgcctcaatc.....attattgtgataaaaaa 2687

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:

1: gb_ba:*

2: gb_hg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_dr:*

10: gb_ro:*

11: gb_srs:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_srs:*

28: em_un:*

29: em_vl:*

30: em_hg_hum:*

31: em_hg_in:*

32: em_hg_ocher:*

33: em_hg_mus:*

34: em_hg_pln:*

35: em_hg_rtd:*

36: em_hg_mam:*

37: em_hg_vit:*

38: em_sy:*

39: em_hgo_hum:*

40: em_hgo_mus:*

41: em_hgo_ocher:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-----------|--------------------|
| 1 | 2685.4 | 99.9 | 2687 | AF494338 | AF494338 Arabidops |
| 2 | 2685.4 | 99.9 | 97263 | AC013258 | AC013258 Arabidops |
| 3 | 1088.8 | 40.5 | 1476 | AY065049 | AY065049 Arabidops |
| 4 | 1068.2 | 39.8 | 1501 | AY087257 | AY087257 Arabidops |
| 5 | 742.8 | 27.6 | 1011 | AY093747 | AY093747 Arabidops |
| 6 | 711.4 | 26.5 | 120877 | AC025808 | AC025808 Genomic s |
| 7 | 534 | 19.9 | 1340 | AY065041 | AY065041 Arabidops |
| 8 | 532.4 | 19.8 | 1359 | AY086340 | AY086340 Arabidops |
| 9 | 530.8 | 19.8 | 1257 | AF134217 | AF134217 Arabidops |
| 10 | 501.4 | 18.7 | 945 | AF372937 | AF372937 Arabidops |
| 11 | 326.4 | 12.1 | 411 | AY074829 | AY074829 Arabidops |
| 12 | 194.8 | 7.2 | 1494 | AF395901 | AF395901 Lycopersi |
| 13 | 156 | 5.8 | 19854 | ATAP22 | 299708 Arabidops |
| 14 | 156 | 5.8 | 198780 | ATCHR1V86 | AL161590 Arabidops |
| 15 | 152.6 | 5.7 | 84196 | ATT3A5 | AL133979 Arabidops |
| 16 | 152.6 | 5.7 | 89469 | ATF18B3 | AL049862 Arabidops |
| 17 | 149 | 5.5 | 831 | AX050690 | AX050690 Sequence |
| 18 | 149 | 5.5 | 831 | BT006310 | BT006310 Arabidops |
| 19 | 149 | 5.5 | 1273 | BT002452 | BT002452 Arabidops |
| 20 | 142 | 5.3 | 1859 | AY050384 | AY050384 Arabidops |
| 21 | 141.4 | 5.3 | 957 | AY097357 | AY097357 Arabidops |
| 22 | 118 | 4.4 | 90425 | F9K20 | AC005679 Arabidops |
| 23 | 117 | 4.4 | 978 | AY090331 | AY090331 Arabidops |
| 24 | 117 | 4.4 | 1515 | AY050430 | AY050430 Arabidops |
| 25 | 113.2 | 4.2 | 11911 | ATRI3C5 | AL021711 Arabidops |
| 26 | 113.2 | 4.2 | 19075 | ATCHR1V49 | AL161545 Arabidops |
| 27 | 110.8 | 4.1 | 855 | AX651445 | AX651445 Sequence |
| 28 | 110.8 | 4.1 | 1217 | AY088379 | AY088379 Arabidops |
| 29 | 110.8 | 4.1 | 1258 | AK118850 | AK118850 Arabidops |
| 30 | 110.6 | 4.1 | 14480 | AP004276 | AP004276 Oryza sat |
| 31 | 108.2 | 3.9 | 1897 | AX653881 | AX653881 Sequence |
| 32 | 104 | 3.9 | 158815 | AP003105 | AP003105 Oryza sat |
| 33 | 96 | 3.6 | 166126 | AP003686 | AP003686 Oryza sat |
| 34 | 93.4 | 3.5 | 1116 | AX653341 | AX653341 Sequence |
| 35 | 93.2 | 3.5 | 192110 | AP004070 | AP004070 Oryza sat |
| 36 | 92.4 | 3.4 | 1374 | AY200612 | AY200612 Arabidops |
| 37 | 85 | 3.2 | 176734 | AP003682 | AP003682 Oryza sat |
| 38 | 77.6 | 2.9 | 2070 | BT006482 | BT006482 Arabidops |
| 39 | 77.6 | 2.9 | 2272 | AK117140 | AK117140 Arabidops |
| 40 | 77.6 | 2.9 | 62052 | AB020744 | AB020744 Arabidops |
| 41 | 71.4 | 2.7 | 349980 | AX344560 | AX344560 Sequence |
| 42 | 70 | 2.6 | 6145 | AX345874 | AX345874 Sequence |
| 43 | 69.6 | 2.6 | 5499 | AX458455 | AX458455 Sequence |
| 44 | 67 | 2.5 | 109135 | AP003408 | AP003408 Oryza sat |
| 45 | 67 | 2.5 | 155955 | AP003288 | AP003288 Oryza sat |

ALIGNMENTS

RESULT 1
LOCUS AF494338 2687 bp DNA linear PLN 20-MAY-2002
DEFINITION Arabidopsis thaliana BZRI protein (BZRI) gene, complete cds.
ACCESSION AF494338
VERSION AF494338.1 GI:20270970
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS Wang,Z.Y., Nakano,T., Gendron,J., He,J., Chen,M., Vafeados,D.,
Yang,Y., Fujikawa,S., Yoshida,S., Asemi,T. and Chory,J.

TITLE Nuclear-localized BZR1 mediates brassinosteroid-induced growth and feedback suppression of brassinosteroid biosynthesis
JOURNAL Dev. Cell 2 (4), 505-513 (2002)
MEDLINE 21966115
PubMed 11970900
REFERENCE 2 (bases 1 to 2687)
AUTHORS Wang, Z.-Y. and Chory, J.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2002) Plant Biology, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Rd., La Jolla, CA 92037, USA

FEATURES
source location/Qualifiers
1..2687
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/map="between ngall1 and ACATPase"
/note="ecotype: Columbia"
1..2687
/gene="BZR1"
1..880
/gene="BZR1"
join(881..1369,1714..2687)
/gene="BZR1"
/product="BZR1 protein"
join(1103..1369,1714..2457)
/gene="BZR1"
/codon_start=1
/product="BZR1 protein"
/protein_id="AA18490.1"
/db_xref="GI:20270971"
/translation="MTSDGATSTSAIAAAAAAAAAARRKPSWREHNNRRRRRAVA
AKIYTGRAQGDVNLPHCDNNEVYLKALCYGAVVEDGTTTKGCPLEGEIAGTS
SRVTYSSONOSPLISAFOSEIPSYOVSPPSSPSPSPSPSPSPSPSPSPSPSPSP
IPSELPSELRISNSCPVTPVSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
VSAPASPTHRHQFTPATPTPEICDESSTUDSGHISQKFAQQQPSASAVPTSPTE
NLVPPAQMSPTNPAFOEIGSSSEFFKENSQVPRWGERIHVDGMDELTLGNGRA
RG"
2147
/gene="BZR1"
/note="compared to BZR1-1D mutant"
/replace="L"
BASE COUNT 781 a 554 c 494 g 858 t

ORIGIN
Query Match 99.9%; Score 2685.4; DB 8; Length 2687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2686; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTTTAGTTTGCCTAATTCATCGAACCTCTGATTCATTCGAAATGTTCCCAACTGC 60
DB 1 ACTTTAGTTTGCCTAATTCATCGAACCTCTGATTCATTCGAAATGTTCCCAACTGC 60
QY 61 GTTGATGTTGGGTTCTCTGCTTTTAAGACATTAGTTACATGCTCTGCACTTTTAC 120
DB 61 GTTGATGTTGGGTTCTCTGCTTTTAAGACATTAGTTACATGCTCTCTGCACTTTTAC 120
QY 121 AAAATAAAGCTCATATCCAAATAATTAACATGATCATACATATATGCGCGGAAC 180
DB 121 AAAATAAAGCTCATATCCAAATAATTAACATGATCATACATATATGCGCGGAAC 180
QY 181 TTGTTATGGACAACCTGTAACCCCTTTTCTTTATGTTCAATGAATATCAAGT 240
DB 181 TTGTTATGGACAACCTGTAACCCCTTTTCTTTATGTTCAATGAATATCAAGT 240
QY 241 TTGTTATGGACAACCTGTAACCCCTTTTCTTTATGTTCAATGAATATCAAGT 300
DB 241 TTGTTATGGACAACCTGTAACCCCTTTTCTTTATGTTCAATGAATATCAAGT 300
QY 301 GATCTAAAGCTTAAATCAATGTAACCAACTAAATACCTTTATAGAACATGTAAT 360
DB 301 GATCTAAAGCTTAAATCAATGTAACCAACTAAATACCTTTATAGAACATGTAAT 360

QY 361 GGTATTCACCAATCTTTATATCATTTGTGAAGTACGAAGAGTAAAAAAGAGAGAGC 420
DB 361 GGTATTCACCAATCTTTATATCATTTGTGAAGTACGAAGAGTAAAAAAGAGAGAGC 420
QY 421 CAGTGTACATACACTAATCAGGACAAAGAGTCAAAATGTTTCTAAAGTGAATTTG 480
DB 421 CAGTGTACATACACTAATCAGGACAAAGAGTCAAAATGTTTCTAAAGTGAATTTG 480
QY 481 TATCAAGAAAAAGTATATTTTATTAATGAATATATCATATGATGTTATATCAAGG 540
DB 481 TATCAAGAAAAAGTATATTTTATTAATGAATATATCATATGATGTTATATCAAGG 540
QY 541 CTTCACGTATATAATTAATTTAAAAACCAATCAAAATGTTGTTTTCGTACACA 600
DB 541 CTTCACGTATATAATTAATTTAAAAACCAATCAAAATGTTGTTTTCGTACACA 600
QY 601 AGTAAAGGACCCACATGAAGAAAGGATCCACGTGTCTCTCTCTCTCTCTCTGTA 660
DB 601 AGTAAAGGACCCACATGAAGAAAGGATCCACGTGTCTCTCTCTCTCTCTCTGTA 660
QY 661 TTATTTGTCATTACTCATTTTATCATATCTCACAGAAAAAAGATTAACATTAAC 720
DB 661 TTATTTGTCATTACTCATTTTATCATATCTCACAGAAAAAAGATTAACATTAAC 720
QY 721 ACAGTTACTAAGGAGTATCTCTGACCTTAACATACCTCTTATATTCACCTCA 780
DB 721 ACAGTTACTAAGGAGTATCTCTGACCTTAACATACCTCTTATATTCACCTCA 780
QY 781 CGTAATCTCAACCTTCCAAAAACATGATTTTACACGTGAGATCGAATACAGAAACA 840
DB 781 CGTAATCTCAACCTTCCAAAAACATGATTTTACACGTGAGATCGAATACAGAAACA 840
QY 841 TGATTTCTTAATATGAATCAATGACTTGAACACACACAGACCAATTTTACATTAGA 900
DB 841 TGATTTCTTAATATGAATCAATGACTTGAACACACACAGACCAATTTTACATTAGA 900
QY 901 TGAATAAATAATTAATTTGTTGGAAGAAAGAGATTTCTTCTTGATTCAGC 960
DB 901 TGAATAAATAATTAATTTGTTGGAAGAAAGAGATTTCTTCTTGATTCAGC 960
QY 961 GAAGGAAAGGATTTCTCTGAGCACTTCTCACTCTCTCTCTCTCTCTCTCTCAT 1020
DB 961 GAAGGAAAGGATTTCTCTGAGCACTTCTCTCACTCTCTCTCTCTCTCTCTCTCAT 1020
QY 1021 CAGTCTACGTTCAACATCTTTCAACCACTTAATCAAGCTCTCTCCGAAATTTGAG 1080
DB 1021 CAGTCTACGTTCAACATCTTTCAACCACTTAATCAAGCTCTCTCCGAAATTTGAG 1080
QY 1081 GGGTTGTTGTTGTTTCCGATGACTTGGATGAGCTTACGTCGATCAGACAGCTGC 1140
DB 1081 GGGTTGTTGTTGTTTCCGATGACTTGGATGAGCTTACGTCGATCAGACAGCTGC 1140
QY 1141 AGCTGCGCGCGGACACAGCGGAGAGAGAACCCGTCGTGAGAGAAAGGAAATTA 1200
DB 1141 AGCTGCGCGCGGACACAGCGGAGAGAGAACCCGTCGTGAGAGAAAGGAAATTA 1200
QY 1201 TCGGAGAGAGAAAGAGAGAGAGAGCTGTAGCTGCGAAGATATCACTGGGCTTAGAGC 1260
DB 1201 TCGGAGAGAGAGAAAGAGAGAGAGAGCTGTAGCTGCGAAGATATCACTGGGCTTAGAGC 1260
QY 1261 TCAAGGTGATTAATTTGCTTAACATTTGATTAATTAAGTCTTTAAAGCTTTTG 1320
DB 1261 TCAAGGTGATTAATTTGCTTAACATTTGATTAATTAAGTCTTTAAAGCTTTTG 1320
QY 1321 TGTGAAGCTGTGTTGGTTGTTGAAGAAAGATGTAATTAATGCAAGGTGAACATTT 1380
DB 1321 TGTGAAGCTGTGTTGGTTGTTGAAGAAAGATGTAATTAATGCAAGGTGAACATTT 1380
QY 1381 CTCCATTTTTCAGATCTGAGCTGTTTATATGTTTGTGATGTTTGAATCTGAAT 1440
DB 1381 CTCCATTTTTCAGATCTGAGCTGTTTATATGTTTGTGATGTTTGAATCTGAAT 1440

1441 CGTTGATTTCAATTTGGTAAATGAGTGAATCTGAGAAATTTGAGGGTTTCTCAAG 1500
 1441 CGTTGATTTCAATTTGGTAAATGAGTGAATCTGAGAAATTTGAGGGTTTCTCAAG 1500
 1501 TGAATTTGATTCATCAGAAACTATGATGATGATGATGATGATGATGATGATGATGAT 1560
 1501 TGAATTTGATTCATCAGAAACTATGATGATGATGATGATGATGATGATGATGATGAT 1560
 1561 TTTCTTTCTAATTTAGAGTATTTAGAGTATTTAGAGTATTTAGAGTATTTAGAGTAT 1620
 1561 TTTCTTTCTAATTTAGAGTATTTAGAGTATTTAGAGTATTTAGAGTATTTAGAGTAT 1620
 1621 CTGTTGTCAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 1621 CTGTTGTCAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 1681 TTTGTTATTTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1740
 1681 TTTGTTATTTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1740
 1741 GCTGGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
 1741 GCTGGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
 1801 GCTTTTCAAGTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 1801 GCTTTTCAAGTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 1861 CCTTTCGCGGATGAAACCAATATGATGATGATGATGATGATGATGATGATGATGAT 1920
 1861 CCTTTCGCGGATGAAACCAATATGATGATGATGATGATGATGATGATGATGATGAT 1920
 1921 GGTGGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1980
 1921 GGTGGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1980
 1981 CCGGTCTCATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 1981 CCGGTCTCATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 2041 AAGCAATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
 2041 AAGCAATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
 2101 GTTTCGACCTGCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2160
 2101 GTTTCGACCTGCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2160
 2161 GAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
 2161 GAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
 2221 TTTTGCACAAACAGACCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2280
 2221 TTTTGCACAAACAGACCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2280
 2281 GTGAACCTGCGGCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
 2281 GTGAACCTGCGGCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
 2341 AGCTCTGAGTTTAAATTTGAGAAATGAGAAATTTGAGAAATTTGAGAAATTTGAG 2400
 2341 AGCTCTGAGTTTAAATTTGAGAAATGAGAAATTTGAGAAATTTGAGAAATTTGAG 2400
 2401 GATGTGGGATGAGAGATTTAGCTTAACTTGAAGAGAGAGAGAGAGAGAGAGAGAT 2460
 2401 GATGTGGGATGAGAGATTTAGCTTAACTTGAAGAGAGAGAGAGAGAGAGAGAGAT 2460
 2461 AAGCAATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
 2461 AAGCAATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
 2521 TATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2580

Db 2521 TATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2580
 Oy 2581 TATTCGAATATGCTGCTCCCAAGCCTTCTCAATGAGAACTTGGAGTGTAT 2640
 Db 2581 TATTCGAATATGCTGCTCCCAAGCCTTCTCAATGAGAACTTGGAGTGTAT 2640
 Oy 2641 TCAATTCGAATATGCTGCTCCCAAGCCTTCTCAATGAGAACTTGGAGTGTAT 2687
 Db 2641 TCAATTCGAATATGCTGCTCCCAAGCCTTCTCAATGAGAACTTGGAGTGTAT 2687

RESULT 2

AC013258/c 97263 bp DNA linear PLN 19-JAN-2001

LOCUS Arabidopsis thaliana chromosome 1 BAC F9E10 genomic sequence.

DEFINITION complete sequence.

ACCESSION AC013258

VERSION AC013258.5 GI:12323880

KEYWORDS HMG.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana (thale cress); Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 97263) Lin, X., Kaul, S., Town, C.D., Benito, M.-I., Creasy, T.H., Haas, B.J., Wu, D., Maiti, R., Romm, C.M., Koo, H., Fujii, C.Y., Uterback, T.R., Barnstead, M.B., Bowman, C.L., White, O., Nierman, W.C., and Frazer, C.M. Arabidopsis thaliana chromosome 1 BAC F9E10 genomic sequence Unpublished

2 (bases 1 to 97263) Lin, X. and Kaul, S.

REFERENCE Direct Submission Submitted (05-NOV-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA. xlin@tigr.org

3 (bases 1 to 97263) Medical Center Dr, Rockville, MD 20850, USA. xlin@tigr.org

REFERENCE Town, C.D. and Kaul, S.

AUTHORS Direct Submission Submitted (19-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA. ctown@tigr.org

TITLE Medical Center Dr, Rockville, MD 20850, USA. ctown@tigr.org

JOURNAL On Jan 19, 2001, this sequence version replaced gi:12280837.

COMMENT Address all correspondence to: atet@tigr.org

BAC clone F9E10 is from Arabidopsis thaliana chromosome 1. The orientation of the sequence is from Sp6 to T7 end of the BAC clone.

Genes were identified by a combination of several methods: Gene prediction programs including GenScan+ (Chris Burge, <http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHM (Mark Borodovsky, <http://genemark.biology.gatech.edu/GeneMark/>), Glimmer (a variant of Glimmer, see Mihaela Pertea, <http://www.tigr.org/softlab/glimmer.htm>), and GeneSplicer (Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org), searches of the complete sequence against a peptide database and the plant EST database at TIGR (<http://www.tigr.org/tdb/cgi.shtml>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

FEATURES

source 1..97263
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Columbia"
 /db_xref="taxon:3702"

Db 361 AGCTAAGGTGATTAATTTGCTTAAACATTTGATTAATTAAGTCTTAAACTCT 420
 Oy 1318 TTGTGTGAAGCTGTTGGTTGTGAAGAGTGTACTTATTCGCAAGGTGAAGAC 1377
 Db 421 TTGTGTGAAGCTGTTGGTTGTGAAGAGTGTACTTATTCGCA----- 470
 Oy 1378 TTTCTCAATTTTTCAGATCTGAGCTGTTTATTGATGTTTGTGATTTGAATCTGA 1437
 Db 471 ----- 470
 Oy 1438 ATTGTGTAATTCATTTGTTAAAGGTTGAATCTGAGATTTGAGGTTTCTCA 1497
 Db 471 ----- 470
 Oy 1498 AAGTAATTTGAATCATCAGAACTATGATGATCTGATTTCTCAAGTAATTAATGG 1557
 Db 471 ----- 470
 Oy 1558 GTTTTCTTCTAATTTTAGATTATTATGTTATGCTAAAGTCTTAATCTTTATGTTAG 1617
 Db 471 ----- 470
 Oy 1618 ATACTGTGTCAGAAATTCATTCATTTGTTTCTTTGCTTACCTGTGATGATGTT 1677
 Db 471 ----- 470
 Oy 1678 TGATGTTATTTGTTTGTCTTTTGTGAGTATCAGGATGCAAGCTTTACCTGTGAG 1737
 Db 471 ----- AGGATGCAAGCTTTTACCTGTGAG 496
 Oy 1738 ATAGCTGGGACTTCATCTCGAGTATCTCATATTCATCAGAACGAGAGCCCTTTCA 1797
 Db 497 ATAGCTGGGACTTCATCTCGAGTATCTCATATTCATCAGAACGAGAGCCCTTTCA 556
 Oy 1798 TCAGCTTTAAAGTCCCATCCCATCTTCAAGTGAAGCCGCTTCTCATCATTCGCG 1857
 Db 557 TCAGCTTTAAAGTCCCATCCCATCTTCAAGTGAAGCCGCTTCTCATCATTCGCG 616
 Oy 1858 AGTCTTCTCGCGGTGAACCAATATAGAACATGTCTCTCAATTCCTTCCCTTCTGAGA 1917
 Db 617 AGTCTTCTCGCGGTGAACCAATATAGAACATGTCTCTCAATTCCTTCCCTTCTGAGA 676
 Oy 1918 AATGTGGCAATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1977
 Db 677 AATGTGGCAATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 736
 Oy 1978 CCACGGTCTCATGCGGACTTCTTAAAGAACCCGAAACCGTTGCTTAATCGGAATTCATC 2037
 Db 737 CCACGGTCTCATGCGGACTTCTTAAAGAACCCGAAACCGTTGCTTAATCGGAATTCATC 796
 Oy 2038 GCTAAGCAATCATGAGGCAATGCTAATCAATCAATGAGGCTTTTAAATTAATCTTCTTCT 2097
 Db 797 GCTAAGCAATCATGAGGCAATGCTAATCAATCAATGAGGCTTTTAAATTAATCTTCTTCT 856
 Oy 2098 GCGGTTTCTGACCTGCTAATGCTGACATGATGCGCAAGTTCATACCTGAGCTACTATA 2157
 Db 857 GCGGTTTCTGACCTGCTAATGCTGACATGATGCGCAAGTTCATACCTGAGCTACTATA 916
 Oy 2158 CTTGAATGATGATGCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2217
 Db 917 CTTGAATGATGATGCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 976
 Oy 2218 AAGTTTGACAACAACAGCAATTCCTGCTTATGAGTGCACCACTCTTCTTCTTCTTCTTCT 2277
 Db 977 AAGTTTGACAACAACAGCAATTCCTGCTTATGAGTGCACCACTCTTCTTCTTCTTCTTCT 1036
 Oy 2278 CTTTGAAACCTGCGCTCAGAGATGCTCTCAATATCTGCTGCTTCTTCTTCTTCTTCTTCT 2337
 Db 1037 CTTTGAAACCTGCGCTCAGAGATGCTCTCAATATCTGCTGCTTCTTCTTCTTCTTCTTCT 1096
 Oy 2338 CAAAGCTGTAGTTTAAATTTGAGAAATGCGCAAGTTAAACCTTGGGAAGAGAGAGATA 2397
 Db 1097 CAAAGCTGTAGTTTAAATTTGAGAAATGCGCAAGTTAAACCTTGGGAAGAGAGAGATA 1156

Oy 2398 CATGATGTGGGATGAGAGATCTTGAGCTTACATTTGAAATGGAAGGCTCGTGTGA 2457
 Db 1157 CATGATGTGGGATGAGAGATCTTGAGCTTACATTTGAAATGGAAGGCTCGTGTGA 1216
 Oy 2458 CATTAACAATAAGCAACCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2517
 Db 1217 CATTAACAATAAGCAACCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1276
 Oy 2518 GAGTATTTTCTTCTTCTGCTCAGATATTTGATCTTTATGGAATCTCATATGTTCTTACT 2577
 Db 1277 GAGTATTTTCTTCTTCTGCTCAGATATTTGATCTTTATGGAATCTCATATGTTCTTACT 1336
 Oy 2578 TATTAACCAATATGCTGCGCAAGGCTTCTCAGATGAGATCTTGAAGTGTGATGATGAT 2637
 Db 1337 TATTAACCAATATGCTGCGCAAGGCTTCTCAGATGAGATCTTGAAGTGTGATGATGAT 1396
 Oy 2638 TATTAACCAATATGCTGCGCAAGGCTTCTCAGATGAGATCTTGAAGTGTGATGATGAT 2697
 Db 1397 TATTAACCAATATGCTGCGCAAGGCTTCTCAGATGAGATCTTGAAGTGTGATGATGAT 1446

RESULT 4
 LOCUS AY087257 1501 bp mRNA linear PLN 14-APR-2003
 DEFINITION Arabidopsis thaliana clone 33367 mRNA, complete sequence.
 ACCESSION AY087257
 VERSION AY087257.1 GI:21405961
 KEYWORDS FLI, CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE
 AUTHORS Haas, B.J., Volkovskiy, N., Town, C.D., Troupkhan, M., Alexandrov, N., Feldmann, K.A., Flavell, R.B., Whits, O. and Salberg, S.L.
 1 (bases 1 to 1501)
 2 (bases 1 to 1501)

TITLE Full-length messenger RNA sequences greatly improve genome annotation
 JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
 MEDLINE 22088475
 PUBMED 12093376

REFERENCE
 AUTHORS Brover, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
 Full-length cDNA from Arabidopsis thaliana

TITLE Unpublished
 JOURNAL 3 (bases 1 to 1501)
 AUTHORS Brover, V., Troupkhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.

COMMENT Direct Submission
 Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made available to RIKR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or later ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genes carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES
 source 1..1501
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"

```

/db xref="taxon:3702"
/clone="33367"
187..1197
/codon_start=1
/product="unknown"
/protein_id="AA064812.1"
/db xref="GI:21592862"
/translation="MTSDGATSTGAAAAAARRKPSRERENRRRRRAVA
AKIYTGRAQGDYNI,PKHCDNNEVLKALCVAGWVEEDGTYRKGCKPLBGEIAGTS
SRVTPYSSONSPLSAFAQSPIDPSYOVSPSPSPRSGPNNMSSTPPRIANGS
IPSLRLRLINSNCPLTPPVNSPTSKPKPLPNMBSIAKOSMAIAKOSMASFNPTA
VSAPASPTHRHQFTPATIPEDCSDSTVDSGHWISFQKFAQOQPSASVPTSPF
NLVPAPOQMSPTNAFAOEIGQSSSEPFENSQVPMWGERIHVDMEDLETLGNRA
RG"
BASE COUNT      389 a      349 c      312 g      451 t
ORIGIN
Query Match      39.8% Score 1068.2; DB 8; Length 1501;
Best Local Similarity 80.4%; Pred. No. 8,7e-216;
Matches 1424; Conservative 0; Mismatches 3; Indels 344; Gaps 1;

Oy 917 ATTTGGAGAGAGAGAGAGATTTCTTCTTCGATTCAGAGAGAGAGAGAGATTT 976
Db 1 ATTTGGAGAGAGAGAGAGATTTCTTCTTCGATTCAGAGAGAGAGAGAGATTT 60

Oy 977 CCTCGAGAGAGAGAGAGAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1036
Db 61 CCTCGAGAGAGAGAGAGAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 120

Oy 1037 AATCTTCAACCACTTATTCAGAGAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1096
Db 121 AATCTTCAACCACTTATTCAGAGAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 180

Oy 1097 TTCCGATGACTTCGATGAGAGAGAGAGAGATTTCTTCTTCTTCTTCTTCTTCTTCTT 1156
Db 181 TTCCGATGACTTCGATGAGAGAGAGAGAGATTTCTTCTTCTTCTTCTTCTTCTTCTT 240

Oy 1157 GCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1216
Db 241 GCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

Oy 1217 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1276
Db 301 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

Oy 1277 TTGCGTAACTTGTATATATATATATATATATATATATATATATATATATATAT 1336
Db 361 TTGCGTAACTTGTATATATATATATATATATATATATATATATATATATATAT 420

Oy 1337 GTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1396
Db 421 GTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 451

Oy 1397 TGTGAGCTTGTATATATATATATATATATATATATATATATATATATATAT 1456
Db 452 TGTGAGCTTGTATATATATATATATATATATATATATATATATATATATAT 451

Oy 1457 GGTAAATGGGTTGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1516
Db 452 GGTAAATGGGTTGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 451

Oy 1517 GAAACTATGATGATCTGATTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1576
Db 452 GAAACTATGATGATCTGATTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 451

Oy 1577 AGTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1636
Db 452 AGTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 451

Oy 1637 TGCATTTGTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1696
Db 452 TGCATTTGTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 451

```

```

Oy 1697 TTTTGTGAGATCAGGAGATCAAGAGCTTTACCGGTGAGATAGCTGGAGATTCATCTC 1756
Db 452 TTTTGTGAGATCAGGAGATCAAGAGCTTTACCGGTGAGATAGCTGGAGATTCATCTC 496

Oy 1757 GAGTAACCTCAATTCATTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1816
Db 497 GAGTAACCTCAATTCATTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 556

Oy 1817 TCCCATCTTACCAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1876
Db 557 TCCCATCTTACCAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616

Oy 1877 CAATTAACAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1936
Db 617 CAATTAACAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 676

Oy 1937 CTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1996
Db 677 CTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 736

Oy 1997 CTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2056
Db 737 CTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 796

Oy 2057 TTGCTAAACAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2116
Db 797 TTGCTAAACAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 856

Oy 2117 GTCCGACACATCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2176
Db 857 GTCCGACACATCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 916

Oy 2177 ACTCTTCACTGTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2236
Db 917 ACTCTTCACTGTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 976

Oy 2237 CATTTCTGCTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2296
Db 977 CATTTCTGCTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1036

Oy 2297 AGCAGATGCTCCAAATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2356
Db 1037 AGCAGATGCTCCAAATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1096

Oy 2357 TTGAGATATGCAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2416
Db 1097 TTGAGATATGCAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1156

Oy 2417 ATCTTGAAGCTTACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2476
Db 1157 ATCTTGAAGCTTACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1216

Oy 2477 CCAATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2536
Db 1217 CCAATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1276

Oy 2537 CAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2596
Db 1277 CAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1336

Oy 2597 CCAGAGCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2656
Db 1337 CCAGAGCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1396

Oy 2657 GTTCAAAAGCATTAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2687
Db 1397 GTTCAAAAGCATTAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1427

```

RESULT 5
 AY093747
 LOCUS
 DEFINITION Arabidopsis thaliana At1g75080/P9E10.7 mRNA, complete cds.

Db 1177 TC 1178

RESULT 8
LOCUS AY086340 1359 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 24157 mRNA, complete sequence.
ACCESSION AY086340
VERSION AY086340.1 GI:21405050
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE Hase, B. J., Volkovskiy, N., Town, C. D., Troukhan, M., Alexandrov, N., Feldmann, K. A., Flavell, R. B., White, O. and Salzberg, S. L.
Full-length messenger RNA sequences greatly improve genome annotation
Genome Biol. 3 (6), RESEARCH0029 (2002)
JOURNAL MEDLINE
PUBMED 12093376
2 (bases 1 to 1359)
REFERENCE Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
Full-length CDNA from Arabidopsis thaliana
Unpublished
3 (bases 1 to 1359)
REFERENCE Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
Direct Submision
Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road, Malibu, CA 90265, USA
This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these CDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genes carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.
Location/Qualifiers
1. 1359
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="24157"
112. 1119
/codon_start=1
/product="unknown"
/protein_id="AA064408.1"
/db_xref="GI:21592457"
/translation="MTSDGATSTSAAMAAAAMATRRKPSWRERENRRRRRRRAVAA
KATYGRAGQNNYLPKHCNNEVYLKALCSAGWVEDGTTTKGKHPLEGDAGSSS
KATYGRSHNOSPSTFSDPILSYVSPSSSPSPRVDGDPNHSITIFPLKNGGIP
SSLPPIRISASAVTPPVSPSPSRNSRSLPTWSPKOSMAAKOSMTSANTPYAV
SADSPTHRQFAPATIPCCDSDSSTUDSGHISQKAAQOCPBSASVPPSPITN
LVAPAPQOLSPNTAAIQEIGSSSEPKENSVKRWBSGRIDVAMELELTLGNGRAH
S"
BASE COUNT 356 a 317 c 285 g 401 t
ORIGIN
Query Match 19.8%; Score 532.4; DB 8; Length 1359;
Best Local Similarity 81.8%; Pred. No. 3.3e-102;

Matches 640; Conservative 0; Mismatches 136; Indels 6; Gaps 2;

QY 1712 AGGATGCAACCTTTACTGGTGAGATACCTGAGCTTCACTGAGTAACCTCCATAT 1771
DB 374 AGGACACAGAGCTCTACTGGTGAGATACCTGAGCTTCACTGAGTAACCTCCATAT 433
QY 1772 CATCAGACAG 1831
DB 434 CTTCCTATACCAAG 493
QY 1832 TTAGCCGCTCTCTTCACTATCTCCGAGTCTCTCGCGGTGAACCAATTAACATGT 1891
DB 494 TCAGCTCTTCT 551
QY 1892 CCTTCACT 1951
DB 552 -CTTCACATCT 610
QY 1952 GAATCTCAACAGTGTGCAAGTACCCGAGTCTCAATGCGGAGCTTCAAGAACCGA 2011
DB 611 GAATCTCAACAGTGTGCAAGTACCCGAGTCTCAATGCGGAGCTTCAAGAACCGA 670
QY 2012 AACCGTGTCTTACCTGGAATCTATGCTTACCAATCCATGGCCAT--TGTAAACAT 2068
DB 671 AACCATGCGCTACTTGGGAATCTTACCAACATCCATGCTGAGCTGCTAAACGT 730
QY 2068 CAATGCGCTCTTAAATATCTCTCTATGCGGCTTCTGAGCTGAGTCTGAGCAGATC 2128
DB 731 CAATGCTCTTGAATCAACCTGCTTATGCGGCTTCTGAGCTGAGTCTGAGCAGATC 790
QY 2128 GCCACAGTGTATACCTGCTACTATCTGATGATGATGATGATGATGATGATGATGAT 2188
DB 791 ATGCGCAGTGTATACCTGCTACTATCTGATGATGATGATGATGATGATGATGATGAT 850
QY 2188 TTGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2248
DB 851 TTGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 910
QY 2248 CTATGCTGCAACCT 2308
DB 911 CTATGCTGCAACCT 970
QY 2309 CAAATCTGCTGCT 2368
DB 971 CAAATCTGCTGCT 1030
QY 2369 AAGTTAAACCTGGAAG 2428
DB 1031 AAGTTAAACCTGGAAG 1090
QY 2429 CACTTGAATGGAAG 2488
DB 1091 CGCTTGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 1150
QY 2489 TC 2490
DB 1151 TC 1152

RESULT 9
AF134217 1257 bp mRNA linear PLN 01-JAN-2000
LOCUS AF134217 Arabidopsis thaliana 107 protein (107) mRNA, complete cds.
DEFINITION AF134217
ACCESSION AF134217
VERSION AF134217.1 GI:6651068
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1257)
AUTHORS Kwon, H.-B.

this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES

Location/Qualifiers

1..945
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAF109-13-019 (P12325)"
/notes="ecotype: Columbia"
1..288
5'UTR
CDS
289..699
/note="unknown protein"
/product="Atg19350/F18014_4"
/protein_id="AKS0077.1"
/db_xref="GI:13937167"
/translation="MSMAKOSMTSLNYPYVASAPSPTHRHQFAPATIPEDCEDSD
STVDSGHMISFQKFAOQPFASMPSTPTNLYKRPAPOLSPNTAAIOEIGOSSEF
KFNISQVKPWEGRHIDVAMEDELTLGNKAS"

3'UTR
BASE COUNT 245 a 232 c 172 g 296 t
ORIGIN

Query Match 18.7% Score 501.4; DB 8; Length 945;
Best Local Similarity 82.0%; Pred. No. 1.2e-95;
Matches 603; Conservative 0; Mismatches 126; Indels 6; Gaps 2;

1759 GTAACATCATATTCATACAGAACAGAGCCCTTTTCATCAGCCTTCAAGTCCATC 1818
1 GCACCTCTTACTCTCCCAACCAAGCTCTTTCTTCCACTTTGATAGCCCATC 60
1819 CCATCTTACCAAGTAGCCCGCTCTTTCATCATTCGCCAGTCTTCCGCGTGAACA 1878
61 TTAATCTTACCAAGTAGCTCTTCTCTTCAATCCCGAGTCTTCTCGAGTTGGAT 120
1879 AATAACACATGTCTCTACATCTTCCCTCTCGAAGTGGGATTCCTTCTCT 1938
121 CCACACATAT---CTCCACATCTTCTCTCTCTCGAAGTGGGATTCCTTCTCT 177
1939 CTTCCTTCCCTCAGAACTTCAACAGTTGTCAAGTACCCACCGGTCTCATGCCGACT 1998
178 CTTCCTCAGCTTGAATCTCAACAGTGTCTCTGTCACTCAACAGTGTATCCCAACT 237
1999 TCTAAGAACCCGAACCGTGTCTTACTGGGAATCTATGCTTAAGCAATCATG3CCAT- 2057
238 TCTAAGAACCCGAACCGTGTCTTACTGGGAATCTTTCACAAACATCATGTCATG 297
2058 --TGTAAACAAATCATGGGCTCTTTTAATATCTTCTCTATGCGGTTTCTGACCTGCT 2115
298 GCTGCTAAACAGTCAATGACTCTTGTGAACCTCCGTTTATGCGGTCTGACCTGCG 357
2116 AGTCCGACATGCGCACACAGCTTTCATACCTGGCTACTATACCTGAATGTAGTGTCT 2175
358 AGTCCGACATGCTGCGCACAGCTTTCATACCTGGCTACTATACCTGAATGTAGTGTCT 417
2176 GACTCTTCACTGTGTGATCTGTGCTATGTGATGAATAGTTTCAAGATTGACAAACAG 2235
418 GACTCTTCACTGTGTGATCTGTGCTATGTGATGAATAGTTTCAAGATTGACAAACAG 477
2236 CATTCTTCCGCTCTATGTGCGCAACCTCTCTCACTTCAATCTTGTGAACCTGGCGCT 2295
478 CATTCTTCCGCTCTATGTGCGCAACCTCTCTCACTTCAATCTTGTGAACCTGGCA 537
2296 CAGCAGATGTCTCAAAATCTGCTGCTTCCAAAGATTTGCTGAAGTTTAA 2355
538 CAGCAGATGTCTCAAAATCTGCTGCTTCCAAAGATTTGCTGAAGTTTAA 597
2356 TTTGAGAAATGAGCAAGTTAAACCTGGGAAGAGAGATATCATGTGTGGATGAG 2415
598 TTTGAGAAATGAGCAAGTTAAACCTGGGAAGAGAGATATCATGTGTGGATGAG 657
2416 GATCTTGAAGTTACATTTGGAATGGAGGCTCTGTGTGACATTAACAATAGGCA 2475

Db 658 GATCTAGACTGACGCTTGAATGGTAAAGCTCATAGTGAATGAATCATGATC 717
Qy 2476 CCCAATGCGATCTC 2490
Db 718 CTGTATGTCATCTC 732

RESULT 11

LOCUS AY074829 411 bp mRNA linear PLN 04-FEB-2002
DEFINITION Arabidopsis thaliana Atg19350/F18014_4 mRNA, complete cds.
ACCESSION AY074829
VERSION AY074829.1 GI:18491116
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophytes;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 411)
Cheuk, R., Chen, H., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L.,
Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K.,
Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A.
and Ecker, J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 411)
Cheuk, R., Chen, H., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L.,
Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K.,
Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A.
and Ecker, J.R.
Direct Submision
Submitted (14-JAN-2002) Salk Institute Genomic Analysis Laboratory
(SIGAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

TITLE

JOURNAL
Arabidopsis full-length cDNA (') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

COMMENT

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H.,
Kim, C.J., Meyers, M.C., Shinn, P., Banh, J., Bowser, L., Dale, J.M.,
Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B.,
Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S.,
Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W.,
Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES

Location/Qualifiers

1..411
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U12325"
/note="This clone is in pUNI 51
ecotype: Columbia"
1..411
CDS

```
/note="unknown protein"
/codon_start=1
/product="AF1919350/F18014_4"
/protein_id="AF169527.1"
/db_xref="GI:18491117"
/translation="MSMAKQSMSTSLNTPVAVSPASPTNRPQHPAPATPECDSD
SSTVSGHMSIFQKFAQOQPFSSAVPTSPFNLVKPAQQLSPMTAAIQEIGQSEF
KFENSQVTPWBEGRIDHVAEMDELTLTGKNAHS"
BASE COUNT      108 a      104 c      89 g      110 t
ORIGIN
Query Match      12.1%; Score 326.4; DB 8; Length 411;
Beet Local Similarity 88.5%; Pred. No. 1.58-58;
Matches 354; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 2058 TGCTAAACATCAATGAGCGCTTTTAATATCTCTTCTATGCGGTTTTCACCTGCTTG 2117
DB 12 TGCTAAACATCAATGAGCTTTTGAACCTGTTTATGCGGTCTGCACTGCCAG 71

QY 2118 TCCGACATCGCCACGAGTTTCATACCTGCTACTATACCTGAATGATGAGTTGA 2177
DB 72 TCTTCTATCATCTCCGAGTTTCATGCTCGGCTCTATACCTGAATGATGAGTTGA 131

QY 2178 CTCTTCACTGTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2237
DB 132 CTCTTCACTGTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 191

QY 2238 ATTCTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2297
DB 192 ATTCTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 251

QY 2298 GCAGATGCTCCAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2357
DB 252 GCATTTGCTCCAAACAGCAGCAATCCAGAGATGCTCAAACTCCGAGTTTAACT 311

QY 2358 TGAGATAGCCAAAGTTAAACCTGAGAGAGAGAGATACATGATGATGATGATG 2417
DB 312 TGAGATAGCCAAAGTTAAAGCCATGGAGAGAGAGATGATGATGATGATGATG 371

QY 2418 TCTTGAAGCTTACATTTGAAATGGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 2457
DB 372 TCTTGAAGCTTACATTTGAAATGGATGATGATGATGATGATGATGATGATG 411

RESULT 12
AF395901      1494 bp      mRNA      linear      PLN 08-JUL-2001
LOCUS      Lycopersicon esculentum mature anther-specific protein LAT61 mRNA,
DEFINITION      complete cds.
ACCESSION      AF395901
VERSION      AF395901.1 GI:14626760
KEYWORDS
SOURCE      Lycopersicon esculentum (tomato)
ORGANISM      Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
Asteridae; Jamidae; Solanales; Solanales; Solanales; Lycopersicon.
1 (bases 1 to 1494)
REFERENCE      McCormick, S.
AUTHORS      McCormick, S.
TITLE      Tomato mature anther specific protein (LAT61).
JOURNAL      Unpublished
PUBDATE      2 (bases 1 to 1494)
AUTHORS      McCormick, S.
TITLE      Direct Submision
JOURNAL      Submitted (27-JUN-2001) Plant Gene Expression Center,
USDA/ARS/UC-Berkeley, 800 Buchanan St., Albany, CA 94710, USA
FEATURES
SOURCE      Location/Qualifiers
1. 1494
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
21..1022
/note="similar to protein encoded by GenBank accession
```

```
number AC025808"
/codon_start=1
/product="mature anther-specific protein LAT61"
/protein_id="AAK1662.1"
/db_xref="GI:14626761"
/translation="MMWEGSPASSRAGAGAGGAGAGVGPESGGGGGGRKPSW
RRNNRRRRRRRAVAAKITGAPAAQGNVLRKCDNNNEVKAICTEAGWIVPECT
TYRGCKPTPMEIGSTNITPSSRRHSPSSRYASPTPSOPSTSSPSSPAD
AMHSHPSFLQNVPSLPPLRISNAPVTPPLSPKPTKPTKPTKPTKPTKPTKPTK
NIPFASAPASPTKVPKPTKPTKPTKPTKPTKPTKPTKPTKPTKPTKPTKPTKPTK
VKP/POPRPMDMTDKGKIDFENFENSVAMWBEGRIDHVAEMDELTLTGKNAH1"
BASE COUNT      428 a      329 c      302 g      435 t
ORIGIN
Query Match      7.2%; Score 194.8; DB 8; Length 1494;
Beet Local Similarity 60.9%; Pred. No. 1.2e-30;
Matches 454; Conservative 0; Mismatches 232; Indels 60; Gaps 6;

QY 1712 AGGATGCAAGCCTTTTACCTGATGATGATGATGATGATGATGATGATGATGATG 1771
DB 337 AGGATGCAAGCCTTTTACCTGATGATGATGATGATGATGATGATGATGATGATG 396

QY 1772 CATCAGAGACAGAGCCCTCTTCATGAGCTTTCAAGTCCATCCATCTTACCAAG 1831
DB 397 CTTACGCGCAGACAGAGCTCCCATCATCATCTTGTCTAGCCCATTTCCATCTTAC 456

QY 1832 TTAGCCGCTTTCTTCAATCATTTCCGAGTCTTCTCGGCTGACCAATACATG 1891
DB 457 CAGTCCAGCTCTCTTCTTCCGAGTCTTCTCGGCTGACCAATACATG 510

QY 1892 CCTCAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1951
DB 511 TATCATATCATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 567

QY 1952 GAATCTCAAGATGTTCTCACTTACCCAGGCTCTCATGCGGCTCTTCAAGACCGA 2011
DB 568 GAATCTCAAGATGTTCTCACTTACCCAGGCTCTCTTCTTCTTCTTCTTCTTCT 627

QY 2012 AACGTTGCTTACCTGAGATCTATGCTTACCAATCCATGCGCTTCTTCAATCA 2071
DB 628 AACCTTT-----CAATTTAGAAATTTGGCAAAAGATCA 663

QY 2072 TGGCGCTTTTATATATCTTTCTATGCGGTTTCTGACCTGCTATGCGACATG 2131
DB 664 TGTTCCTTTTAAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720

QY 2132 ACCAGTTTCATACCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2191
DB 721 TTTAGCGTTTATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 780

QY 2192 ATTCTGCTATGATTAAGCTTTTCAAGATTTTCAAGATTTTCAAGATTTTCT 2251
DB 781 ATTCAAGCGCATGATTAAGCTTTTCAAGATTTTCAAGATTTTCAAGATTTTCT 819

QY 2252 TGGTCCCAACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2311
DB 820 ATGTTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 877

QY 2312 ATATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2371
DB 878 -TATGATATATATACAGACAGAGGATTAAGATTAAGATTAAGATTAAGATTA 936

QY 2372 TTAACCTTGGAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 2431
DB 937 TCAAGCATGAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 996

QY 2432 TTGGAATAGGAGAGCTTCTGCTTGA 2457
DB 997 TTGGAATAGGAGAGCTTCTGCTTGA 1022

RESULT 13
ATAP22
```


| | | | |
|---|-------|--|--------|
| D | 35380 | ACGAGGTTCTTAAAGCTCTCTGCTCGAAGCTGGATCGTCGAACGATGGCACC | 354339 |
| O | 1359 | CTTATCGCAAGGTGA-----GACTTCTCCATTTTTTCCAGATCTGAGCTT | 1405 |
| D | 35440 | CTTATCGCAAGTAAATCGAAATTCGGGATTTCATTTTAGATCGGAATCTGAGCTT | 35499 |
| O | 1406 | -----GTTTATGATGTTTGTGATGTTGAATCTGAATTCGTTGATTT | 1449 |
| D | 35500 | CTCTCAATTACTCATTAATTCATTCGATGTAATCTGGCCGATGTGTGTTTAAATCAATC | 35559 |
| O | 1450 | CAATTGGGTTAAAT-----GGGTGATCTGAGAAATTTGAGGGTTTTC | 1497 |
| D | 35560 | GAATTGTAGAGCTTTGAGATCTTGGCATTTGGATTTCACTTTTCACTGTCGATTTGG | 35619 |
| O | 1498 | AACTGATTTGATTCATCAGAAACTAGATGGAATCTGATTTTCTCAAGTGAATTTATGG | 1557 |
| D | 35620 | GTTGAATTTTCAATTTCCCCCTAATTTGATGAATACATGCTGTGTGTGATTTAGAACT | 35679 |
| O | 1558 | GTT-----TCTTTCTAATTTTGAAGTTAATTTGGATGCTAAAGTCTTAAATC | 1606 |
| D | 35680 | GATTATGAATTCGATGGTTTCTGCTTTGGATCATTAATTCATGATCATTAAGTTTG | 35739 |
| O | 1607 | TTTTATGATGAACTTGGTCCAAAGTCAATGCAATGGTTTCTTTGCTTACCTGAT | 1666 |
| D | 35740 | ATTATTTAGAAATTTGGGAAATCTCTTTGCTTCTTGAGCTTGACTTAAACATTGAC | 35799 |
| O | 1667 | TGATGATGTTGATTTGTTATTTGTTGCTTTGT----- | 1702 |
| D | 35800 | TTTGAAGGTGATCATCATGTTGTTTATTTAGTCTAGTCTGTGATCTAAATGGTT | 35859 |
| O | 1703 | -----TGAAGTACAGGATGCAAGCCTTTTACCTGATAGATAGCTGGACTTCATCTCG | 1757 |
| D | 35860 | GATTTCTTGGGAAACAGGGGTTTACGACACACATCAATATTTAGAAATCTTACAA | 35919 |
| O | 1758 | AGTAATCTCATTTATCAACAACAAGCCCTCTTTCATGAGCTTTCAAGTCCAT | 1817 |
| D | 35920 | CTTACACCAAAATTCATCATCAACCAAGTCCAAATCATCAAGTTTTCAGAGTCTCGC | 35979 |
| O | 1818 | CCCATTTACCAAGTTAGCCGCTCTTCTTCATCAATCCCGAGTCCCTCTCGGGGAAAC | 1877 |
| D | 35980 | ACCTTGATCAACGAAATCGATCTCATCTCTCCGAGTCCATCTCGTATGACGG | 36039 |
| O | 1878 | AAA--TAAACAATGCTCTTACATTC--TTCCCTTCTCAGAAATGGTGGCATTCCTC | 1934 |
| D | 36040 | AAACCTTCTTCAATCTCTTCTTCCTGCTTACAAACATCGGTTCTTGATTTCTCGC | 36099 |
| O | 1935 | TTCTCTTCTTCCCTCAGAAATCTCAACAAGTTGTCAAGTACCCACCGGCTCATCGCC | 1994 |
| D | 36100 | TAACTTTCACCTCTTAAATATTCMAACAGTGGCGCTGTGACTCTCGTTGTCACTCC | 36159 |
| O | 1995 | GACTTTTAAGAACCGGAAACGTTGCTCTTACCTGGAATCTATCGGTAAAGCAATCATGGC | 2054 |
| D | 36160 | TACTTCTGTGTTGAAGCGGAAATTAATCTTCAAGCAATTAACAA----- | 36207 |
| O | 2055 | CATTGCTTAAACATGATGGCGCTTTTAAATATCTTTCTATGCGGTTTTCACCTGCG | 2114 |
| D | 36208 | -----TGGCGGTCTTAAACAGTTTGAACATCGGTTTTCGTATCTGTGCTCATC | 36261 |
| O | 2115 | TAGTCGACACATCGCCACCACTTTCATACCTGGCTATCTATCTGAATGTATGATC | 2174 |
| D | 36262 | TA---GACTCTGTGCTGTGATCAACAAACGCACTTCAATACCGGATGTGATGATC | 36317 |
| O | 2175 | ---TGACTTTCACATGTTATCTGTGATCTTGGATTAAGCTTTCAAGAGTTTGCACAA | 2231 |
| D | 36318 | CGAGAGATTCGATCGAGGATTCAGGAAGGTGATCAATTTCCATCTACTGCTCTAC | 36377 |
| O | 2232 | ACAGCATTCTCTGCTCTATGTGTCACCACTCTCTCACTTCAATCTTGTGAACCTGC | 2291 |
| D | 36378 | TTTACCAATTTAACTTTTCAAGCAATCTTCTATGCGCATGATAT----- | 36425 |
| O | 2292 | GCCTCAGACATGTCTCAAAATCTGCTCTTCCAGAGATGTCTAAAGCTGTGATT | 2351 |
| D | 36426 | -----GAAAGGTGAGCTGGGAGATGTCCGAAATGAATGGGAGGGGAGAGTT | 36476 |

| | | | |
|----|-------|--|-------|
| D | 21778 | CTTCAATGCCCCCGGGAGGAGAGAGAGAGAGATCATTCGTCCGGACGTACTCCGA | 21719 |
| OY | 1179 | CGTAGAGAAAAGGAGAAATATTCGAGAGAGAAAAGCGAGAAAGACTGTAGTCGGA | 1238 |
| D | 21718 | CGTAGAAGAAGAGAGAGAACATAAGAGAGAGAGAGAGAGAGAGAACCATCCTGCTTA | 21655 |
| OY | 1239 | AGATTATACACTGGGCTTAGAGCTCAGAGTGATTTAAATTGCCCTAAACATTGTGATTA | 1298 |
| D | 21658 | AGATTACTCTGGCTTAGAGCTCAGAGTAACCTTAAGCTTCCTTAAGCACCTGCATPACA | 21599 |
| OY | 1299 | ATGAAGTCCCTAAAGCTCTTGTGTGAAGCTGTGGGTGTTGAGAGAGATGTACTTA | 1358 |
| D | 21598 | ACGAGGCTTTAAAGCTCTGCTCGAACCTGGTGGATCGTAGAGACATGGCACCA | 21539 |
| OY | 1359 | CTATGCAAGGTGAA-----GACTTCTCAATTTTTCCAGATCTGAGCTT | 1405 |
| D | 21538 | CTTATGCGCAAGGTAAAGTGAATTCGGAATTTCAAATTTTAGATCGGAATCTGAGCTT | 21479 |
| OY | 1406 | -----GTTTATGATGTTTTGATGTTGAACTGAACTGAACTGTTGATTT | 1449 |
| D | 21478 | CTCTCAATTACTCATATTCAATCGATGATCTGTGGAGTTGTGTGTTTAAATTCATC | 21419 |
| OY | 1450 | CATTGTGTTAAAT-----GGGTTTGAATCTGAGAAATTTGAGSGTTTCTCA | 1497 |
| D | 21418 | GAAATGTTAGAGACTCTTGAGATCTTCGATTTGGATTTCACTTTCACTTCGATTTGG | 21359 |
| OY | 1496 | AAGTGAATTTGAATCATCAGAACTATGATGATCGATCTGATTTCTCAAGTAAATTTATGG | 1557 |
| D | 21358 | GTTGAATTTCAATTTCCCCTATTGGATTTGAATCTGGCTCTTGATGATTAAGAACT | 21299 |
| OY | 1558 | GTT-----TCTTTCTAATTTTAGAGTTATATTGTTAGTATGCTAAAGCTTAATC | 1606 |
| D | 21298 | GATTATGAATTCGATGCTTTTCGCTTTGGATCATCTAATTCATAGATCATGTAAAGTTTG | 21239 |
| OY | 1607 | TTTTATGATGATCTGTCCTCAAGCTCATGCAATGTGTGTTCTTTTCTTACCCTGTGAT | 1666 |
| D | 21238 | ATTATTTAGAAAAATTTGGGGAATCTCTTTGCTTCTTGAGCTTGACTAAACATTGAC | 21179 |
| OY | 1667 | TGATTGATGTTGATGTTATTTGTTTGTGTTTGT----- | 1702 |
| D | 21178 | TTTGAGGTTGCATCATCTGTTGTTTGAATTTAGTCTAGCTCTTGTCACCTAAATGTT | 21119 |
| OY | 1703 | -----TGAGTATCAGGGATGCAAGCTTTTACCTGGTGAATAGCTGGGACTTCATCTCG | 1757 |
| D | 21118 | GATTCTTGAGAAACAGGGGTTTAGCACACAGATCAGAAATTTACAGAACTCCTACAA | 21059 |
| OY | 1758 | AGTAACTCATTTTATCATCAGAACAGAGCCCTCTTTCATCAGCCTTTCAMAGTCCAT | 1817 |
| D | 21058 | CTTCAGACAAATTTATCATCAATCCAACCAAGTCCACATCATCAGCTTTTCCAAGTCTGC | 20999 |
| OY | 1818 | CCCATCTTACCAAGTTAGCCCGCTCTTTCATCATTTCCGAGTCTTCTCGGGTGAAC | 1877 |
| D | 20998 | ACCTTCTGACAGAGAGTCCAGTCTCATCATCTTCCGAGTCCATCTGCTATGACGG | 20939 |
| OY | 1878 | AAA--TAAACATGTCCTCTCAATTC--TTCCCTTCTCAGAAATGTTGGATTCCTTC | 1934 |
| D | 20938 | AAACCTTCTTCATACCTTCTTCCGCTTCTTCCACCAACATAGCTTCTTCGATTTCTGC | 20879 |
| OY | 1935 | TTCTCTTCCCTCCAGATCTCAAAACGTTGTCAGTTAACCCACCGGCTTCATGCGC | 1994 |
| D | 20878 | TAACTTCCACTCTTAGAATATCCAAACAGTGGCGCTGAGCTCTCCGTTGCATCTCC | 20819 |
| OY | 1995 | GACTTCTAAGAACCGGAAACGTTGCTTAACTGGGAATCTATGCTTAAGCATCATGCG | 2054 |
| D | 20818 | TACTTCTGTGTTGAAAGCGGAACTTACTTCAGAGCAATTAACAA----- | 20771 |
| OY | 2055 | CATTGCTAAACATCATGCGCTCTTTTAATATCTCTTCTATGCGGTTTCTGACCTGC | 2114 |
| D | 20770 | -----TGGCGGGCTTTTACAGCTTTTAGAGATCCGCTTTTCCGATCTCTGCTCCATC | 20717 |
| OY | 2115 | TAGTCCAGACATGCGCACAGTTTCAATCCCTGGCTACTAATCCTGAATGTGATGATC | 2174 |
| D | 20716 | TA---GACTGTGTGTGTGTCACCAAGCGCACTTAACAAATCCGGAATGTGATAGTTC | 20661 |

| | | | | | |
|--------------|-------|----------------|--|---|-----------------|
| Y | 2175 | --- | tgactcttccac | ctgttgattcttgagcatttgatataagcttttcagaaagttgcacacaca | 2231 |
| Db | 20660 | CGAGAGGAGATTTC | GCATTCGAGAGTTCA | GGAGAGGTGATTCATATTTCCAAATCTACTGCTCTAC | 20601 |
| Y | 2232 | ACAGCCATTCTCT | CGCTCTTATATGATGCCAACCTCTTACCTTCATCAATCTTGAAACCTGC | 2291 | |
| Db | 20600 | TTCAACCAACATTT | TAACCTTGTTTCAGCAAACTTTATGGCCATTGATAT | ----- | 20553 |
| Y | 2292 | GCCTCAGCAGATG | TCTCCAAATACTGCTGCTTTCACAGATTTGGTCAAAGCTCTGAGTT | 2351 | |
| Db | 20552 | ----- | GAAAGGTCTGAGCTGGGGAATGTCGGGAATGATGGAGAGGGGCAAGTT | 20502 | |
| Y | 2352 | TAAATTTGAGAA | TATGCGCAAGTTAAACCTGGGAAAGAGAGAGATACATGATGGGAT | 2411 | |
| Db | 20501 | TGAGTTTGAGAA | TGGAAACAGTTTAAGCCCTGGGAAAGCAAAATGATTCATGAAATGGGCT | 20442 | |
| Y | 2412 | GGAGATCTTGAG | CTTACACTTGGAAATGGGAAGCGCTGTGATTACATA | 2461 | |
| Db | 20441 | AGAAATCTCGAA | CTCACTCTCGGTGGCACTAAAGCCAGATGCTGAATAA | 20392 | |
| RESULT 15 | | | | | |
| ATT3A5/c | | | 84196 bp | DNA | linear |
| LOCUS | | | | | PLN 26-JAN-2000 |
| DEFINITION | | | Arabidopsis thaliana DNA chromosome 3, BAC clone T3A5. | | |
| ACCESSION | | | AL132979 | | |
| VERSION | | | AL132979.2 | | GI:6782244 |
| KEYWORDS | | | | | |
| SOURCE | | | | | |
| ORGANISM | | | Arabidopsis thaliana (chale crese), | | |
| | | | Arabidopsis thaliana | | |
| | | | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| | | | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | |
| | | | rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. | | |
| REFERENCE | | | 1 | | |
| AUTHORS | | | Bleecker,H., Mewes,H.W., Lemcke,K., Mayer,K.F.X., Quetier,F. and Salanoubat,M. | | |
| JOURNAL | | | Unpublished | | |
| REFERENCE | | | 2 (bases 1 to 84196) | | |
| AUTHORS | | | EU Arabidopsis sequencing project. | | |
| TITLE | | | Direct Submission | | |
| JOURNAL | | | Submitted (25-JAN-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mpi-biochem.mpg.de, mayer@mpi-biochem.mpg.de Project Coordinator: Marcel Salanoubat and Francis Quetier, Groupe ment d'Interet Public, Centre National de Sequençage - GENOSCOPE, 2 rue Gaston Cremieux, BP191, 91006 Evry Cedex, France; http://www.genoscope.cns.fr | | |
| COMMENT | | | On Jan 27, 2000 this sequence version replaced gi:6434247. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/. Location/Qualifiers | | |
| FEATURES | | | 1..84196 | | |
| SOURCE | | | /organism="Arabidopsis thaliana" | | |
| | | | /mol_type="genomic DNA" | | |
| | | | /variety="Columbia" | | |
| | | | /db_xref="taxon:3702" | | |
| | | | /chromosome="3" | | |
| | | | 2983..3338 | | |
| LTR | | | /note="355 bp LTR" | | |
| | | | 3339..7564 | | |
| misc_feature | | | /note="LTR-terminoson" | | |
| LTR | | | 7565..7922 | | |
| | | | /note="367bp LTR" | | |
| gene | | | 12339..13197 | | |
| | | | /gene="T3A5.10" | | |
| CDS | | | join(12339..12665,12742..12803,12888..13055,13125..13197 | | |
| | | | /gene="T3A5.10" | | |
| | | | /note="similarity to cyclin-dependent kinase inhibitor (ICK1), Arabidopsis thaliana, EMBL:AT094772_1" | | |
| | | | /codon_start=1 | | |
| | | | /product="putative protein" | | |

/protein_id="CAB62432.1"
 /db_xref="GI:6561966"
 /translation="MAVRRRRRRDVEENGVTITTVKRRKMEEVLDVSRITILPCV
 QATNRGJITVARKSAGSETSVVITVRRRDSPIVEEQCIIEEDSSISCCSTSEKSR
 IEEVDLEENNGDRELETSMITVDLNRKSESNMSSVAIVEDSRRLRSLRSLHETV
 KEHELDPFOVAEKDLRNKLRLCSEMKYNFDFPEKDLPGGRYEWVKLPN"
 12339. 12665
 /gene="T3A5.10"
 /number=1
 12666. 12741
 /gene="T3A5.10"
 /number=1
 12742. 12803
 /gene="T3A5.10"
 /number=2
 12804. 12887
 /gene="T3A5.10"
 /number=2
 12888. 13055
 /gene="T3A5.10"
 /number=3
 13056. 13124
 /gene="T3A5.10"
 /number=3
 13125. 13197
 /gene="T3A5.10"
 /number=4
 13588. 16470
 /gene="T3A5.20"
 complement(join(15858. 15901,16014. 16470))
 /gene="T3A5.20"
 complement(join(15858. 15901,16014. 16470))
 /gene="T3A5.20"
 /note="similarity to predicted proteins, Arabidopsis
 thaliana"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB62433.1"
 /db_xref="GI:6561967"
 /translation="MARLEHNSNGGAHRTSFSENEVELRSEKSKNSNNIRSSFS
 MPSADRFAPSTVDYSMIPADEIRPLKGLIPKETSHTRTLGEELLTTEEGSGWDGNTF
 SLRRIILSSSFSTKGTRELLGLKRTYRSKTKDKVNEEVLSDHKHISGNVATREC
 QVADTR"
 complement(15858. 15901)
 /gene="T3A5.20"
 /number=1
 complement(15902. 16013)
 /gene="T3A5.20"
 /number=1
 complement(16014. 16470)
 /gene="T3A5.20"
 /number=2
 18303. 19931
 /gene="T3A5.30"
 complement(18303. 19931)
 /gene="T3A5.30"
 complement(18303. 19931)
 /gene="T3A5.30"
 /codon_start=1
 /product="gcarecrow-1 like 7 (SCL7)"
 /protein_id="CAB62434.1"
 /db_xref="GI:6561968"
 /translation="MAVYCTDSGNLMAIAOQLIKQKQOQSHOQOEEOEPNPNR
 PSFGTLTSGSGFSDPQVINDPGFPHLBEHONAAVASEFSDDEMESTLINDGAS
 QNTDFPPIYGHDPVPSFSLGAPSYLRVVKDQASAOQLPPASTAIVSPSPSPQ
 HPPRPPPPDLDLNPQIRKAIHDYARKETEDTLIRIKESVSGDPIQVGYPAE
 ALSHKETSPSSSSSLDEPILSYKTLINDACPSKFAHLTANOAILLEATONSNIH
 VDFGIFQIQWSALQALATRRSGKPTRIRISGIPASISDPSPPSILACNRLRDR
 AILDLNEFVEYVLTPIQLNGSSFRVDEVLVNFMLELYKLDELTTATYKLRILA
 RSLMRITLTLEVEYSLNRVEFARVKNQSLNFTYSAVESLEPNLDROSKELRVERV
 FGRRIMDLVREDDDNNKPGTRFGLMEEKQWRVLMERKAGFEPVPSNVAVASOAKLLLM
 NYNVTSLVSESGFSLAMNVPLLTIVSSNR"
 complement(18303. 19931)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 19:11:09 ; Search time 683 Seconds
(without alignments)
10619.898 Million cell updates/sec

Title: US-09-995-938A-2

Perfect score: 2687
1 attcattcttgcctcaatc.....attcattctgagataaaaaa 2687

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-----------|----------------------|
| 1 | 1068.2 | 39.8 | 1501 | AAC37477 | Arabidopsis thaliana |
| 2 | 532.4 | 19.8 | 1359 | AAC37137 | Arabidopsis thaliana |
| 3 | 532.4 | 19.8 | 1418 | AAC36495 | Arabidopsis thaliana |
| 4 | 149 | 5.5 | 831 | ABZ12580 | Arabidopsis thaliana |
| 5 | 141.6 | 5.3 | 642 | AAAC42317 | Arabidopsis thaliana |
| 6 | 121.4 | 4.5 | 413 | ABX22943 | Human GDP-mannose |
| 7 | 110.8 | 4.1 | 1213 | AAAC51043 | Arabidopsis thaliana |
| 8 | 110.8 | 4.1 | 1217 | AAAC42550 | Arabidopsis thaliana |

| | | | | | | |
|----|-------|-----|--------|----|----------|---------------------|
| 9 | 109.6 | 4.1 | 284 | 25 | ABX32572 | Human GDP-mannose |
| 10 | 93.2 | 3.5 | 459 | 22 | AAH87878 | Peppermint plant o |
| 11 | 6145 | 2.6 | 6145 | 24 | ABL32972 | Human immune syste |
| 12 | 69.6 | 2.6 | 5499 | 24 | ABO65971 | Human angiogenesis |
| 13 | 65 | 2.4 | 47108 | 24 | ABL31510 | Signal transductio |
| 14 | 62.6 | 2.3 | 9642 | 24 | ABL32357 | Human immune syste |
| 15 | 61.6 | 2.3 | 6143 | 25 | ABZ10194 | Haematopoietic cel |
| 16 | 60.8 | 2.3 | 16033 | 24 | ABL33404 | Human immune syste |
| 17 | 60.6 | 2.3 | 5678 | 24 | ABL33138 | Human immune syste |
| 18 | 60.4 | 2.2 | 6070 | 24 | ABO67130 | Human angiogenesis |
| 19 | 60.4 | 2.2 | 6070 | 24 | ABL70372 | Chemically treated |
| 20 | 60.4 | 2.2 | 6070 | 24 | ABL33679 | Human immune syste |
| 21 | 60.4 | 2.2 | 6070 | 24 | ABL34579 | Human metastasis a |
| 22 | 60.4 | 2.2 | 6954 | 24 | ABL33391 | Human immune syste |
| 23 | 60.4 | 2.2 | 14006 | 24 | ABL33958 | Human immune syste |
| 24 | 60.2 | 2.2 | 5430 | 22 | AA546292 | Tumour suppressor |
| 25 | 60.2 | 2.2 | 6283 | 24 | ABL32088 | Human immune syste |
| 26 | 60.2 | 2.2 | 113515 | 24 | ABL34174 | Human cervical can |
| 27 | 60 | 2.2 | 612 | 22 | AAH71471 | Human immune syste |
| 28 | 60 | 2.2 | 15373 | 24 | ABL32466 | Tumour suppressor |
| 29 | 59.6 | 2.2 | 5649 | 22 | AA546384 | Human chemically p |
| 30 | 59.6 | 2.2 | 5649 | 24 | ABK40008 | Human immune syste |
| 31 | 59.6 | 2.2 | 5649 | 24 | ABL32849 | Human chemically m |
| 32 | 59.6 | 2.2 | 16891 | 24 | ABN80059 | Human immune syste |
| 33 | 59.6 | 2.2 | 16891 | 24 | ABL32653 | Human metastasis a |
| 34 | 59.6 | 2.2 | 16891 | 24 | ABL34505 | Human EST associa |
| 35 | 59.4 | 2.2 | 529 | 25 | ABX52180 | Human metastasis a |
| 36 | 59.4 | 2.2 | 20486 | 24 | ABL34611 | Human metastasis a |
| 37 | 59.4 | 2.2 | 34769 | 22 | AA546774 | Tumour suppressor |
| 38 | 59.2 | 2.2 | 580 | 24 | ABQ34864 | Oligonucleotide fo |
| 39 | 59.2 | 2.2 | 580 | 24 | ABQ34865 | Human immune syste |
| 40 | 59.2 | 2.2 | 6112 | 24 | ABL32489 | Bovine immune syste |
| 41 | 59 | 2.2 | 446 | 25 | ABL38235 | Bovine immune syste |
| 42 | 59 | 2.2 | 5195 | 24 | ABL32921 | Human immune syste |
| 43 | 59 | 2.2 | 6351 | 24 | ABL33747 | Human immune syste |
| 44 | 59 | 2.2 | 6944 | 24 | ABK34027 | Human DNA for stag |
| 45 | 59 | 2.2 | 9539 | 22 | AA545347 | Chemically pretrea |

ALIGNMENTS

RESULT 1

ID AAC37477 standard; DNA: 1501 BP.

AC AAC37477;
DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 17521.

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17521.

XX Hybridisation assay; genetic mapping; gene expression control;

KM protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

OS

XX Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 09-MAR-1999; 99US-0123180.

PR 23-MAR-1999; 99US-0123548.

PR 29-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0126785.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

| | | |
|----|--------------|---------------|
| PR | 16-APR-1999; | 99US-0129845. |
| PR | 19-APR-1999; | 99US-0130077. |
| PR | 21-APR-1999; | 99US-0130049. |
| PR | 23-APR-1999; | 99US-0130510. |
| PR | 28-APR-1999; | 99US-0130891. |
| PR | 30-APR-1999; | 99US-0131449. |
| PR | 30-APR-1999; | 99US-0132048. |
| PR | 04-MAY-1999; | 99US-0132407. |
| PR | 05-MAY-1999; | 99US-0132484. |
| PR | 06-MAY-1999; | 99US-0132485. |
| PR | 06-MAY-1999; | 99US-0132486. |
| PR | 07-MAY-1999; | 99US-0132487. |
| PR | 11-MAY-1999; | 99US-0132863. |
| PR | 14-MAY-1999; | 99US-0134218. |
| PR | 14-MAY-1999; | 99US-0134219. |
| PR | 14-MAY-1999; | 99US-0134221. |
| PR | 18-MAY-1999; | 99US-0134370. |
| PR | 19-MAY-1999; | 99US-0134768. |
| PR | 20-MAY-1999; | 99US-0134941. |
| PR | 21-MAY-1999; | 99US-0135124. |
| PR | 24-MAY-1999; | 99US-0135535. |
| PR | 25-MAY-1999; | 99US-0135629. |
| PR | 27-MAY-1999; | 99US-0136021. |
| PR | 28-MAY-1999; | 99US-0136392. |
| PR | 01-JUN-1999; | 99US-0136782. |
| PR | 03-JUN-1999; | 99US-0137222. |
| PR | 04-JUN-1999; | 99US-0137528. |
| PR | 07-JUN-1999; | 99US-0137702. |
| PR | 08-JUN-1999; | 99US-0137724. |
| PR | 10-JUN-1999; | 99US-0138094. |
| PR | 10-JUN-1999; | 99US-0138540. |
| PR | 14-JUN-1999; | 99US-0138847. |
| PR | 16-JUN-1999; | 99US-0139119. |
| PR | 17-JUN-1999; | 99US-0139452. |
| PR | 18-JUN-1999; | 99US-0139453. |
| PR | 18-JUN-1999; | 99US-0139454. |
| PR | 18-JUN-1999; | 99US-0139455. |
| PR | 18-JUN-1999; | 99US-0139456. |
| PR | 18-JUN-1999; | 99US-0139457. |
| PR | 18-JUN-1999; | 99US-0139458. |
| PR | 18-JUN-1999; | 99US-0139459. |
| PR | 18-JUN-1999; | 99US-0139460. |
| PR | 18-JUN-1999; | 99US-0139461. |
| PR | 18-JUN-1999; | 99US-0139463. |
| PR | 18-JUN-1999; | 99US-0139750. |
| PR | 21-JUN-1999; | 99US-0139763. |
| PR | 22-JUN-1999; | 99US-0139817. |
| PR | 23-JUN-1999; | 99US-0139899. |
| PR | 23-JUN-1999; | 99US-0140353. |
| PR | 24-JUN-1999; | 99US-0140354. |
| PR | 24-JUN-1999; | 99US-0140695. |
| PR | 28-JUN-1999; | 99US-0140823. |
| PR | 29-JUN-1999; | 99US-0140991. |
| PR | 30-JUN-1999; | 99US-0141287. |
| PR | 01-JUL-1999; | 99US-0141842. |
| PR | 01-JUL-1999; | 99US-0142154. |
| PR | 02-JUL-1999; | 99US-0142055. |
| PR | 06-JUL-1999; | 99US-0142390. |
| PR | 08-JUL-1999; | 99US-0142803. |
| PR | 09-JUL-1999; | 99US-0142920. |
| PR | 12-JUL-1999; | 99US-0142977. |
| PR | 13-JUL-1999; | 99US-0143542. |
| PR | 14-JUL-1999; | 99US-0143544. |
| PR | 15-JUL-1999; | 99US-0144005. |
| PR | 16-JUL-1999; | 99US-0144085. |
| PR | 16-JUL-1999; | 99US-0144086. |
| PR | 19-JUL-1999; | 99US-0144325. |
| PR | 19-JUL-1999; | 99US-0144331. |
| PR | 19-JUL-1999; | 99US-0144332. |
| PR | 19-JUL-1999; | 99US-0144333. |
| PR | 19-JUL-1999; | 99US-0144334. |
| PR | 19-JUL-1999; | 99US-0144335. |
| PR | 20-JUL-1999; | 99US-0144632. |
| PR | 20-JUL-1999; | 99US-0144684. |
| PR | 21-JUL-1999; | 99US-0144814. |
| PR | 21-JUL-1999; | 99US-0145086. |
| PR | 21-JUL-1999; | 99US-0145088. |
| PR | 22-JUL-1999; | 99US-0145085. |
| PR | 22-JUL-1999; | 99US-0145087. |
| PR | 22-JUL-1999; | 99US-0145089. |
| PR | 23-JUL-1999; | 99US-0145192. |
| PR | 23-JUL-1999; | 99US-0145145. |
| PR | 23-JUL-1999; | 99US-0145218. |
| PR | 23-JUL-1999; | 99US-0145224. |
| PR | 26-JUL-1999; | 99US-0145276. |
| PR | 27-JUL-1999; | 99US-0145913. |
| PR | 27-JUL-1999; | 99US-0145918. |
| PR | 27-JUL-1999; | 99US-0145919. |
| PR | 28-JUL-1999; | 99US-0145951. |
| PR | 02-AUG-1999; | 99US-0146386. |
| PR | 02-AUG-1999; | 99US-0146388. |
| PR | 02-AUG-1999; | 99US-0146389. |
| PR | 03-AUG-1999; | 99US-0147038. |
| PR | 04-AUG-1999; | 99US-0147204. |
| PR | 04-AUG-1999; | 99US-0147302. |
| PR | 05-AUG-1999; | 99US-0147192. |
| PR | 05-AUG-1999; | 99US-0147260. |
| PR | 06-AUG-1999; | 99US-0147303. |
| PR | 06-AUG-1999; | 99US-0147416. |
| PR | 09-AUG-1999; | 99US-0147493. |
| PR | 09-AUG-1999; | 99US-0147935. |
| PR | 10-AUG-1999; | 99US-0148171. |
| PR | 11-AUG-1999; | 99US-0148319. |
| PR | 12-AUG-1999; | 99US-0148341. |
| PR | 13-AUG-1999; | 99US-0148565. |
| PR | 13-AUG-1999; | 99US-0148684. |
| PR | 16-AUG-1999; | 99US-0149368. |
| PR | 17-AUG-1999; | 99US-0149175. |
| PR | 18-AUG-1999; | 99US-0149426. |
| PR | 20-AUG-1999; | 99US-0149722. |
| PR | 20-AUG-1999; | 99US-0149723. |
| PR | 20-AUG-1999; | 99US-0149929. |
| PR | 23-AUG-1999; | 99US-0149902. |
| PR | 23-AUG-1999; | 99US-0149930. |
| PR | 25-AUG-1999; | 99US-0150566. |
| PR | 26-AUG-1999; | 99US-0150884. |
| PR | 27-AUG-1999; | 99US-0151065. |
| PR | 27-AUG-1999; | 99US-0151066. |
| PR | 27-AUG-1999; | 99US-0151080. |
| PR | 30-AUG-1999; | 99US-0151103. |
| PR | 31-AUG-1999; | 99US-0151138. |
| PR | 01-SEP-1999; | 99US-0151930. |
| PR | 07-SEP-1999; | 99US-0152363. |
| PR | 10-SEP-1999; | 99US-0153070. |
| PR | 13-SEP-1999; | 99US-0153758. |
| PR | 15-SEP-1999; | 99US-0154018. |
| PR | 16-SEP-1999; | 99US-0154779. |
| PR | 20-SEP-1999; | 99US-0154779. |
| PR | 22-SEP-1999; | 99US-0155139. |
| PR | 22-SEP-1999; | 99US-0155486. |
| PR | 24-SEP-1999; | 99US-0155659. |
| PR | 28-SEP-1999; | 99US-0156458. |
| PR | 29-SEP-1999; | 99US-0156596. |
| PR | 04-OCT-1999; | 99US-0157117. |
| PR | 05-OCT-1999; | 99US-0157753. |
| PR | 06-OCT-1999; | 99US-0157865. |
| PR | 07-OCT-1999; | 99US-0158029. |
| PR | 08-OCT-1999; | 99US-0158232. |
| PR | 12-OCT-1999; | 99US-0158369. |
| PR | 13-OCT-1999; | 99US-0159293. |
| PR | 13-OCT-1999; | 99US-0159294. |
| PR | 13-OCT-1999; | 99US-0159295. |

QY 2657 GTTCAAGCATTATTGTAGATTAATAA 2687
|||||
Db 1397 GTTCAAGCATTATTGTAGATTAATAA 1427

RESULT 2
AAC37137
ID AAC37137 standard; DNA; 1359 BP.

XX AAC37137;
AC
XX 17-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 16310.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
PN
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147203.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147453.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

```

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 16-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 19.8%; Score 532.4; DB 21; Length 1359;
 Best Local Similarity 81.8%; Pred. No. 1.2e-118;
 Matches 640; Conservative 0; Mismatches 136; Indels 6; Gaps 2;

```

QY 1712 AGGATGCAAGCTTACTGCTGAGATGCTGGGACTTCATCTGAGTAACCTCATATT 1771
DB 374 AGGACACAAAGCTTACTGCTGAGATGCTGGGACTTCATCTGAGTAACCTCATATT 433
QY 1772 CATCAGAAACGAGCCCTTTTCATGCGCTTTCGAAGTCCCATCTTTCAGAG 1831
DB 434 CTTCCATTAACCAAGCTCTTTCTTCCATTTGATGAGCCCATCTTATCTTACCAAG 493

```

```

QY 1832 TTAGCCCGCTCTCTTCATCATTCGCCGAGTCTTCTCGCGGTAACCAATAACAACATGT 1891
DB 494 TCAGTCTCTCTCTCTTCATTCGCCGAGTCTTCTCGAGTGGATGATCACACATAT-- 551
QY 1892 CCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1951
DB 552 -CTCACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 610
QY 1952 GAATCTCAACAGTGTCCAGTTACCCACCCGCTCTATGCCGACTCTAAGAACCCGA 2011
DB 611 GAATCTCAACAGTGTCCAGTTACCCACCCGCTCTATGCCGACTCTAAGAACCTCA 670
QY 2012 ACCGTCCTTAACCTGGAATCTATCGCTAAGCAATCATGGCCAT--TGTAAACAAT 2068
DB 671 AACCATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 730
QY 2069 CAATGCGCTCTTAAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2128
DB 731 CAATGACTCTCTTGAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 790
QY 2129 GCCACCACTTTCATACCTCTGCTACTATCTCTGATGTGATGATGTGATGTGATGTG 2188
DB 791 ATCCCACTTTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 850
QY 2189 TTGATTCGTCATGTGATTAAGCTTTCAGAGTTGCAACAACAGCATCTCTGCTT 2248
DB 851 TTGATTCGTCATGTGATTAAGCTTTCAGAGTTGCAACAACAGCATCTCTGCTT 910
QY 2249 CTATGTCGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2308
DB 911 CTATGTCGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 970
QY 2309 CAATPACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2368
DB 971 CAATPACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1030
QY 2369 AAGTTAAACCTGGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2428
DB 1031 AAGTTAAACCTGGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1090
QY 2429 CACTTGGAAATGGAAGGCTCTGCTGATGATGATGATGATGATGATGATGATGATGAT 2488
DB 1091 CACTTGGAAATGGAAGGCTCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1150
QY 2489 TC 2490
DB 1151 TC 1152

```

RESULT 3
 AAC36495
 ID AAC36495 standard; DNA; 1418 BP.

AAC36495;

17-OCT-2000 (First entry)

Arabidopsis thaliana DNA fragment SEQ ID NO: 14021.

Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

| | |
|-----------------|---------------|
| PR 05-MAR-1999; | 99US-0123180. |
| PR 09-MAR-1999; | 99US-0123548. |
| PR 23-MAR-1999; | 99US-0125788. |
| PR 25-MAR-1999; | 99US-0126264. |
| PR 29-MAR-1999; | 99US-0126785. |
| PR 01-APR-1999; | 99US-0127462. |
| PR 06-APR-1999; | 99US-0128234. |
| PR 16-APR-1999; | 99US-0128714. |
| PR 19-APR-1999; | 99US-0129845. |
| PR 21-APR-1999; | 99US-0130077. |
| PR 23-APR-1999; | 99US-0130449. |
| PR 28-APR-1999; | 99US-0130510. |
| PR 30-APR-1999; | 99US-0130891. |
| PR 04-MAY-1999; | 99US-0131449. |
| PR 05-MAY-1999; | 99US-0132407. |
| PR 06-MAY-1999; | 99US-0132484. |
| PR 06-MAY-1999; | 99US-0132485. |
| PR 07-MAY-1999; | 99US-0132486. |
| PR 11-MAY-1999; | 99US-0132863. |
| PR 14-MAY-1999; | 99US-0134256. |
| PR 14-MAY-1999; | 99US-0134218. |
| PR 14-MAY-1999; | 99US-0134221. |
| PR 14-MAY-1999; | 99US-0134370. |
| PR 18-MAY-1999; | 99US-0134768. |
| PR 19-MAY-1999; | 99US-0134941. |
| PR 20-MAY-1999; | 99US-0135124. |
| PR 21-MAY-1999; | 99US-0135353. |
| PR 24-MAY-1999; | 99US-0135629. |
| PR 25-MAY-1999; | 99US-0136021. |
| PR 27-MAY-1999; | 99US-0136392. |
| PR 28-MAY-1999; | 99US-0136782. |
| PR 01-JUN-1999; | 99US-0137222. |
| PR 03-JUN-1999; | 99US-0137528. |
| PR 04-JUN-1999; | 99US-0137502. |
| PR 07-JUN-1999; | 99US-0137724. |
| PR 08-JUN-1999; | 99US-0138099. |
| PR 10-JUN-1999; | 99US-0138540. |
| PR 10-JUN-1999; | 99US-0138847. |
| PR 14-JUN-1999; | 99US-0139119. |
| PR 16-JUN-1999; | 99US-0139452. |
| PR 17-JUN-1999; | 99US-0139453. |
| PR 18-JUN-1999; | 99US-0139457. |
| PR 18-JUN-1999; | 99US-0139456. |
| PR 18-JUN-1999; | 99US-0139457. |
| PR 18-JUN-1999; | 99US-0139458. |
| PR 18-JUN-1999; | 99US-0139459. |
| PR 18-JUN-1999; | 99US-0139460. |
| PR 18-JUN-1999; | 99US-0139461. |
| PR 18-JUN-1999; | 99US-0139462. |
| PR 18-JUN-1999; | 99US-0139463. |
| PR 18-JUN-1999; | 99US-0139750. |
| PR 21-JUN-1999; | 99US-0139763. |
| PR 22-JUN-1999; | 99US-0139817. |
| PR 23-JUN-1999; | 99US-0139899. |
| PR 23-JUN-1999; | 99US-0140353. |
| PR 24-JUN-1999; | 99US-0140354. |
| PR 24-JUN-1999; | 99US-0140695. |
| PR 28-JUN-1999; | 99US-0140823. |
| PR 29-JUN-1999; | 99US-0140991. |
| PR 30-JUN-1999; | 99US-0141287. |
| PR 01-JUL-1999; | 99US-0141842. |
| PR 01-JUL-1999; | 99US-0142154. |
| PR 02-JUL-1999; | 99US-0142055. |
| PR 06-JUL-1999; | 99US-0142390. |
| PR 08-JUL-1999; | 99US-0142803. |
| PR 09-JUL-1999; | 99US-0142920. |
| PR 12-JUL-1999; | 99US-0142977. |
| PR 13-JUL-1999; | 99US-0143542. |
| PR 14-JUL-1999; | 99US-0143624. |
| PR 15-JUL-1999; | 99US-0144005. |
| PR 16-JUL-1999; | 99US-0144085. |
| PR 19-JUL-1999; | 99US-0144325. |
| PR 19-JUL-1999; | 99US-0144331. |
| PR 19-JUL-1999; | 99US-0144332. |
| PR 19-JUL-1999; | 99US-0144333. |
| PR 19-JUL-1999; | 99US-0144334. |
| PR 19-JUL-1999; | 99US-0144335. |
| PR 20-JUL-1999; | 99US-0144632. |
| PR 20-JUL-1999; | 99US-0144864. |
| PR 21-JUL-1999; | 99US-0144814. |
| PR 21-JUL-1999; | 99US-0145086. |
| PR 22-JUL-1999; | 99US-0145088. |
| PR 22-JUL-1999; | 99US-0145087. |
| PR 22-JUL-1999; | 99US-0145089. |
| PR 22-JUL-1999; | 99US-0145192. |
| PR 23-JUL-1999; | 99US-0145145. |
| PR 23-JUL-1999; | 99US-0145228. |
| PR 26-JUL-1999; | 99US-0145276. |
| PR 27-JUL-1999; | 99US-0145913. |
| PR 27-JUL-1999; | 99US-0145918. |
| PR 27-JUL-1999; | 99US-0145919. |
| PR 28-JUL-1999; | 99US-0145951. |
| PR 02-AUG-1999; | 99US-0146386. |
| PR 02-AUG-1999; | 99US-0146388. |
| PR 02-AUG-1999; | 99US-0146389. |
| PR 03-AUG-1999; | 99US-0147038. |
| PR 04-AUG-1999; | 99US-0147204. |
| PR 04-AUG-1999; | 99US-0147302. |
| PR 05-AUG-1999; | 99US-0147326. |
| PR 06-AUG-1999; | 99US-0147303. |
| PR 06-AUG-1999; | 99US-0147416. |
| PR 09-AUG-1999; | 99US-0147493. |
| PR 09-AUG-1999; | 99US-0147935. |
| PR 10-AUG-1999; | 99US-0148171. |
| PR 11-AUG-1999; | 99US-0148319. |
| PR 12-AUG-1999; | 99US-0148341. |
| PR 13-AUG-1999; | 99US-0148565. |
| PR 13-AUG-1999; | 99US-0148684. |
| PR 16-AUG-1999; | 99US-0149368. |
| PR 17-AUG-1999; | 99US-0149175. |
| PR 18-AUG-1999; | 99US-0149426. |
| PR 20-AUG-1999; | 99US-0149722. |
| PR 20-AUG-1999; | 99US-0149723. |
| PR 20-AUG-1999; | 99US-0149929. |
| PR 23-AUG-1999; | 99US-0149902. |
| PR 23-AUG-1999; | 99US-0149930. |
| PR 25-AUG-1999; | 99US-0150566. |
| PR 26-AUG-1999; | 99US-0150884. |
| PR 27-AUG-1999; | 99US-0151065. |
| PR 27-AUG-1999; | 99US-0151066. |
| PR 27-AUG-1999; | 99US-0151080. |
| PR 30-AUG-1999; | 99US-0151303. |
| PR 31-AUG-1999; | 99US-0151438. |
| PR 01-SEP-1999; | 99US-0151930. |
| PR 07-SEP-1999; | 99US-0152363. |
| PR 10-SEP-1999; | 99US-0153070. |
| PR 13-SEP-1999; | 99US-0153758. |
| PR 15-SEP-1999; | 99US-0154018. |
| PR 16-SEP-1999; | 99US-0154039. |
| PR 20-SEP-1999; | 99US-0154779. |
| PR 22-SEP-1999; | 99US-0155119. |
| PR 23-SEP-1999; | 99US-0155486. |
| PR 24-SEP-1999; | 99US-0155659. |
| PR 28-SEP-1999; | 99US-0156458. |
| PR 29-SEP-1999; | 99US-0156596. |
| PR 04-OCT-1999; | 99US-0157117. |

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152353.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 5 3%; Score 141.6; UB 21; Length 642;
Best Local Similarity 72.0%; Pred. No. 3,9e+24;
Matches 183; Conservative 1; Mismatches 70; Indels 0; Gaps 0;
QY 1119 CTACGTGACATCAGCAGCTGAGCTGCGCGCCGACAGCGCGGAGAGAAAGCCGT 1178
DB 128 CTTMATGCGCCCGGAGGAGGAGGAGGAGGAGATCATGTCGGACGTACTCGGA 187
QY 1179 CCGTGAAGAAAGGAGAAATATCGGAGAGAGAAAGACGAGAAAGAGCTGTACCTCGA 1238
DB 188 CGTGAAGAGAGAGAGAGAAATATAGAGAGAGAAAGAAAGAAAGAGGACATCATCTGCTA 247
QY 1239 AGATATACACTGGCTTAGAGCTCAGGTGATATATATTGCCCTAAATTTGATATATA 1298
DB 248 AGATTACTCTGCTTTAGAGCTCAAGTACTATAGCTTCTTACGACTGCGATACCA 307
QY 1299 ATGAGTCCCTTAAAGCTTTTGTGTAAGCTGTTGGTGTGTAAGAAAGTGTACTA 1358
DB 308 ACGAGTTCTTAAAGCTTCTGTCTCGAAGCTGTTGGATGTCGAAGACATGGACCA 367
QY 1359 CTTATCGCAGAGTG 1372
DB 368 CTTATCGCAGAGGG 381
RESULT 6
ABX22943
ID ABX22943 standard; cDNA; 413 BP.
XX
AC ABX22943;
XX
DT 10-FEB-2003 (first entry)
XX
DE Human GDP-mannose 4,6-dehydratase (GM4, 6D) DNA #5000.
XX
KW Human: GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
KW complex carbohydrate; gene replacement therapy; immunosuppressive;
KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
KW antiasthmatic; vasotropic.
XX
OS Homo sapiens.
XX
XX US2002110548-A1.
PN 15-AUG-2002.
XX
BD 11-JUN-2001; 2001US-0878574.
PF 22-NOV-1996; 96US-0753233.
XX
XX 03-DEC-1997; 97US-0984246.
PR 09-SEP-1998; 98US-0149674.
PR 14-JUN-1999; 99US-0333177.
XX
XX (GENY) GENETICS INST INC.
XX
XX Sullivan F, Kriz R, Kumar R;
XX
DR WPI; 2003-066673/06.
XX
XX New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
PT peptide, for manufacturing complex carbohydrates, or as targets for
PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant
PT rejection -

| | | |
|----|-------------|---------------|
| PR | 06-APR-1999 | 99US-01287214 |
| PR | 08-APR-1999 | 99US-01382834 |
| PR | 16-APR-1999 | 99US-01298455 |
| PR | 19-APR-1999 | 99US-01300777 |
| PR | 21-APR-1999 | 99US-01304449 |
| PR | 23-APR-1999 | 99US-01305010 |
| PR | 23-APR-1999 | 99US-01305911 |
| PR | 26-APR-1999 | 99US-01312448 |
| PR | 30-APR-1999 | 99US-01320498 |
| PR | 04-MAY-1999 | 99US-01324407 |
| PR | 04-MAY-1999 | 99US-01324454 |
| PR | 06-MAY-1999 | 99US-01324465 |
| PR | 06-MAY-1999 | 99US-01324476 |
| PR | 07-MAY-1999 | 99US-01328853 |
| PR | 11-MAY-1999 | 99US-01342126 |
| PR | 14-MAY-1999 | 99US-01342181 |
| PR | 14-MAY-1999 | 99US-01342119 |
| PR | 14-MAY-1999 | 99US-01343211 |
| PR | 14-MAY-1999 | 99US-01343710 |
| PR | 15-MAY-1999 | 99US-01347468 |
| PR | 20-MAY-1999 | 99US-01349411 |
| PR | 21-MAY-1999 | 99US-01351524 |
| PR | 22-MAY-1999 | 99US-01352533 |
| PR | 25-MAY-1999 | 99US-01356229 |
| PR | 27-MAY-1999 | 99US-01360211 |
| PR | 28-MAY-1999 | 99US-01367822 |
| PR | 01-JUN-1999 | 99US-01372232 |
| PR | 03-JUN-1999 | 99US-01375728 |
| PR | 04-JUN-1999 | 99US-01375602 |
| PR | 07-JUN-1999 | 99US-01377724 |
| PR | 08-JUN-1999 | 99US-01380594 |
| PR | 10-JUN-1999 | 99US-01385400 |
| PR | 10-JUN-1999 | 99US-01388477 |
| PR | 11-JUN-1999 | 99US-01393911 |
| PR | 15-JUN-1999 | 99US-01394322 |
| PR | 16-JUN-1999 | 99US-01394353 |
| PR | 17-JUN-1999 | 99US-01394492 |
| PR | 18-JUN-1999 | 99US-01394654 |
| PR | 18-JUN-1999 | 99US-01394655 |
| PR | 18-JUN-1999 | 99US-01394661 |
| PR | 18-JUN-1999 | 99US-01394662 |
| PR | 18-JUN-1999 | 99US-01394663 |
| PR | 18-JUN-1999 | 99US-01394670 |
| PR | 18-JUN-1999 | 99US-01397635 |
| PR | 21-JUN-1999 | 99US-01396171 |
| PR | 22-JUN-1999 | 99US-01396819 |
| PR | 23-JUN-1999 | 99US-01403553 |
| PR | 24-JUN-1999 | 99US-01403554 |
| PR | 24-JUN-1999 | 99US-01406655 |
| PR | 28-JUN-1999 | 99US-01406823 |
| PR | 29-JUN-1999 | 99US-01409091 |
| PR | 30-JUN-1999 | 99US-01412877 |
| PR | 01-JUL-1999 | 99US-01418442 |
| PR | 01-JUL-1999 | 99US-01421554 |
| PR | 02-JUL-1999 | 99US-01420555 |
| PR | 06-JUL-1999 | 99US-01429050 |
| PR | 08-JUL-1999 | 99US-01428203 |
| PR | 09-JUL-1999 | 99US-01432820 |
| PR | 12-JUL-1999 | 99US-01432820 |
| PR | 13-JUL-1999 | 99US-01432877 |
| PR | 14-JUL-1999 | 99US-01435324 |
| PR | 15-JUL-1999 | 99US-01440055 |
| PR | 16-JUL-1999 | 99US-01440085 |
| PR | 16-JUL-1999 | 99US-01440086 |
| PR | 19-JUL-1999 | 99US-01443325 |
| PR | 19-JUL-1999 | 99US-01443325 |

PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154033.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159838.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 110.8; DB 21; Length 1213;
Best Local Similarity 70.5%; Pred. No. 1.5e+16;
Matches 148; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 1161 CGCGAGAGGAAGCCCTGCGAGAGAAAGGAGATATATCGAGAGAGAAAGCGGA 1220
DB 155 CGCGAGACTAGAACGCCGACTTGGAAGAGAGAAACAACAACGCGAGAGCGCGGA 214
QY 1221 GAAGACCTGTAGCTGCGAAGATATACACTGGCTTAGAGCTCAAGCTGATATATTC 1280
DB 215 GACGAGGANTTGGCGCTAAGATCTTCGACGAGCTTAGAGATTCAGAACTTCAAGCTCC 274
QY 1281 CTAAACATTGTGTATATATGAAGTCTTAAAGCTTTGTGAGAGCTGGTGGTTG 1340
DB 275 CTAAACCTGCGCAACAAAGAGTCTCAAGCTTTATGCAATGAAGCTGGTGGACTG 334
QY 1341 TTGAAGAGATGGACTACTTATCGCAAGG 1370
DB 335 TAGAAGACGACGGAACCTTACCGCAAGG 364

RESULT 8

AAC42550 standard; DNA; 1217 BP.

ID AAC42550;

AC AAC42550;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35978.

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

OS Arabidopsis thaliana.

XX Arabidopsis thaliana.

XX Arabidopsis thaliana.

XX Arabidopsis thaliana.

XX Arabidopsis thaliana.

XX Arabidopsis thaliana.

XX Arabidopsis thaliana.

XX Arabidopsis thaliana.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132486.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135153.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137282.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140921.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 22-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147036.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149724.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.1%; Score 110.8; DB 21; Length 1217;
Best Local Similarity 70.5%; Pred. No. 1.5e-16;
Matches 148; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 1161 CGGCGAGAGAAAGCCGCTGCTGAGAGAAAGGAAATATCGAGAGAGAGAAAGCGGA 1220
DB 155 CGGGGACTTAGACCCGACTTGGAAAGAGAGAAACAAACGCGAGCGGGGAA 214
QY 1221 GAAGAGCTGAGCTGCGGAGATATACACTGGGCTTAGAGCTCAAGTGATATATTTGC 1280
DB 215 GAGAGGCAATGGCGGTAGATCTTCGAGAGACTTAAGATTCATGAAACTTCAAGCTCC 274
QY 1281 CTAACATTTGATATATATAGAGTCCCTTAAGCTTTGTGTTGAAGCTGTTGGTTG 1340
DB 275 CTAACACTGGGAGAACATGATGATCTTCAAGCTTTATGCAATGAGCTGTTGACTG 334
QY 1341 TTGAAGAAGATGTACTACTTATGCGAAG 1370
DB 335 TAGAAGACGACGAACTTATCCGCAAG 364

RESULT 9

ABX32572 standard; cDNA; 284 BP.

ID ABX32572;

AC ABX32572;

DT 11-FEB-2003 (first entry)

DB Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #14629.

XX Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
XX cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
XX arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
XX complex carbohydrate; gene replacement therapy; immunosuppressive;
XX antiinflammatory; antirheumatic; antibacterial; cerebroprotective;
XX antiasthmatic; vasoactive;
XX Homo sapiens.

PN US2002110548-A1.
XX 15-AUG-2002.
PD 11-JUN-2001; 2001US-0878574.
XX 22-NOV-1996; 96US-0753233.
XX 03-DEC-1997; 97US-0984246.
XX 09-SEP-1998; 98US-0149674.
XX 14-JUN-1999; 99US-0333177.
XX (GENY) GENETICS INST INC.
XX Sullivan F, Kriz R, Kumar R;
XX WPI; 2003-066673/06.

PT New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
PT peptide, for manufacturing complex carbohydrates, or as targets for
PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant
PT rejection

XX Disclosure; SEQ ID NO 14631; 6pp; English.

XX The invention relates to a composition comprising a human GDP-mannose
XX 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
XX GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
XX in a mammalian subject and for treating or ameliorating diseases affected
XX by the level of cellular fucosylation or diseases affected by the
XX fucosylation of glycoconjugates. These diseases include arthritis,
XX transplant rejection, asthma, sepsis, reperfusion injury, stroke or
XX infection. The GM4,6D peptide or a polynucleotide encoding it is also
XX useful for manufacturing complex carbohydrates and as targets for
XX screening small molecule antagonists of the activity of the enzyme. The
XX polynucleotide is useful in developing an assay for defects in the
XX enzyme, as well as in gene replacement therapy. Sequences
XX ABX17942-ABX17944 and ABX17947-ABX17948 represent DNA molecules encoding
XX human GM4,6D peptides of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html.

XX Sequence 284 BP; 95 A; 62 C; 89 G; 38 T; 0 other;

Query Match 4.1%; Score 109.6; DB 25; Length 284;
Best Local Similarity 73.3%; Pred. No. 1.5e-16;
Matches 154; Conservative 0; Mismatches 54; Indels 2; Gaps 1;

QY 1163 GCGAGAGAAAGCCGCTGCTGAGAGAAAGGAGATATATCGG--AGAGAGAAAGCGGA 1220
DB 42 GCGGAGAAAGAAAGCCGCTGCTGAGAGAAAGGAGACACAGAAAGAGAAAGCGAA 101
QY 1221 GAAGAGCTGAGCTGCGGAGATATACACTGGGCTTAGAGCTCAAGTGATATATTTGC 1280
DB 102 GAAGAGCAATGACAGCAAAATATATCTCAGACTTCAAGCTCAGGGGAACTTCAACTTGC 161
QY 1281 CTAACATTTGATATATATAGAGTCCCTTAAGCTTTGTGTTGAAGCTGTTGGTTG 1340
DB 162 CAAAGCAATTTGAGACAAACACAGCTTGAAGAACTTCTGCGCAAGACTGTTGGTGG 221
QY 1341 TTGAAGAAGATGTACTACTTATGCGAAG 1370
DB 222 TAGAAGAAAGGAGCACTTATCCGCAAG 251

RESULT 10

AAH87878 standard; cDNA; 459 BP.

ID AAH87878;

AC AAH87878;

DT 25-SEP-2001 (first entry)

DE Peppermint plant oil gland expressed cDNA 234.
 XX
 KW Peppermint plant oil gland cell; terpenoid essential oil; resin;
 KM genetic mapping; antisense suppression; recombinant expression; ss.
 XX
 OS Mentha x piperita.
 XX
 PN W0200153319-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 19-JAN-2001, 2001WO-US02567.
 XX
 PR 20-JAN-2000, 2000US-0177264.
 XX
 PA (CROT/) CROTEAU R. B.
 PA (LANG/) LANGE B. M.
 PA (WILD/) WILDUNG M. R.
 PI Croteau RB, Lange BM, Wildung MR;
 DR WPI, 2001-488706/53.
 XX
 PT New nucleic acid molecules corresponding to mRNA molecules expressed in
 PT peppermint oil glands for enhancing expression of plant oil gland cell
 PT proteins -
 XX
 PS Claim 1, Page 165, 251pp; English.
 XX
 CC The invention relates to nucleic acid molecules (AAH87645-AAH88116) that
 CC correspond to all or part of a mRNA molecule expressed in plant oil
 CC gland cells, especially peppermint and plant oil glands that produce
 CC terpenoid essential oils and resins. The nucleic acids are useful for
 CC genetically mapping a plant genome for genes expressed in plant oil
 CC gland cells and to suppress (for example by antisense suppression) or
 CC enhance their expression (for example by genetically transforming a
 CC plant cell with a replicable expression vector that expresses one or more
 CC proteins naturally expressed in plant oil gland cells). The nucleic acids
 CC are also useful for recombinant expression of plant oil gland proteins
 CC required for terpenoid essential oil and/or resin production in bacterial
 CC and/or yeast cells.
 CC
 XX
 SQ Sequence 459 BP, 106 A, 141 C, 119 G, 93 T, 0 other;
 XX
 Query Match 3.5%; Score 93.2; DB 22; Length 459;
 Best Local Similarity 65.2%; Pred. No. 1.7e-12;
 Matched 137; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
 XX
 QY 1161 CGCGGAGAGAAAGCCGTCGTGAGAGAAAGGAGATATCGAGAGAGAAAGCGCA 1220
 DB 122 CGGCGACGAGCGTCCGACATGAGAGAGCGGAGACACACAGCGAGCGGCGC 181
 QY 1221 GAAAGCTGTAGCTGCGAAGATATACATCGGCTTAGAGCTCAAGGTATATATTTGC 1280
 DB 182 GCGCGCGATCGCGCGAGAGATCTTCGCGCTGAGAGATGACGCACTACAAGCTCC 241
 QY 1281 CTAACATTTGTATATATGAGTCTTAAAGCTTTTGTGTAAGCTGTTGGGTTG 1340
 DB 242 CCAAGGACTGCGCAACACAGAGTCTCAAGGCTCTGCAATGAAAGCTGCTGGGTCA 301
 QY 1341 TTGAAGAGATGATCTACTTATCGCAAG 1370
 DB 302 TCCAGAGAGCGGACCACTTACAGAAAG 331
 XX
 RESULT 11
 ABL32972
 ID ABL32972 standard; DNA; 6145 BP.
 XX
 AC ABL32972;
 XX
 DT 26-MAR-2002 (first entry)
 XX

DE Human immune system associated gene SEQ ID NO: 945.
 XX
 KW Human, immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianemic; cytosolic; noctropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; db.
 XX
 OS Homo sapiens.
 XX
 PN W0200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001, 2001WO-EP07537.
 XX
 PR 30-JUN-2000, 2000DE-1032529.
 PR 01-SEP-2000, 2000DE-1043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 PA
 PI Olek A, Piepenbrock C, Berlin K,
 XX
 DR WPI, 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation.
 XX
 PS Claim 1, SEQ ID NO 945; 32pp + Sequence listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 CC
 XX
 SQ Sequence 6145 BP, 1493 A, 99 C, 1309 G, 3244 T, 0 other;
 XX
 Query Match 2.6%; Score 70; DB 24; Length 6145;
 Best Local Similarity 50.9%; Pred. No. 2.4e-06;
 Matches 166; Conservative 0; Mismatches 160; Indels 0; Gaps 0;
 XX
 QY 1378 TTCTCCATTTTTCAGATCTGAGCTGTTTATGATGTTTGAATCTGA 1437
 DB 5420 TTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5479
 QY 1438 ATTGCGTATTCATTGCGTTAATGCGTTGAATCGAATTTGAGGTTTCTCA 1497
 DB 5480 TTTTGTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 5539
 QY 1498 AAGTAAATTTGATCATCAGAACTATGATGATGATGATTTCTCAAGAGATTATG 1557
 DB 5540 TATGCGTTTATTTTATTTTATTTTATTTTATTTTATCGTTTATTTTATTTT 5599
 QY 1558 GTTTCTTTCTAATTTTATGATTTATTTGATGCTAAAGTCTTAATCTTTATG 1617
 DB 5600 TTTTATTTATTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTT 5659
 QY 1618 ATACTTGTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1677
 DB 5660 TTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 5719
 QY 1678 TGATGATTTATGTTTCTTTGTT 1703
 DB 5720 TTCGTTTTTTTATTTTATTTATTT 5745
 XX

CC sequence of the invention are useful for the diagnosis and therapy of
 CC disease associated with signal transduction e.g. solid tumours and
 CC cancer. ABK3158-ABK31545 represent chemically pretreated genomic DNA
 CC sequences of different genes associated with signal transduction, or
 CC their complementary sequences.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.

SO Sequence 47108 BP; 13283 A; 429 C; 9740 G; 23656 T; 0 other;

Query Match 2.4%; Score 65; DB 24; Length 47108;
 Best Local Similarity 49.7%; Pred. No. 0.0001;
 Matched 222; Conservative 0; Mismatches 220; Indels 5; Gaps 2;

QY 1258 AGCTCAGGATGATTATATATGCTTAACATGCTGATATATGAAGTCTTAAAGCTCT 1317
 DB 195 AGCTAAGAAATTTTATTTTATTTATATATATATATATATATATATATATATAT 254
 QY 1318 TTGCTTGAAGCTGTTGGGTTGTAAGA-AGATGCTACTTATCGCAAGTGAGA 1376
 DB 255 ATGTATATATTTTATTTTATATATATATATATATATATATATATATATATAT 314
 QY 1377 CTTTCTCAATTTTTCAGATCTGAGCTGTTTATGATGTTTGAATGTTGAATCTG 1436
 DB 315 TTTTGAATTTGTTTTTTTATAGTATGATTTAGATTTAGTGAATGTAAGCT 374
 QY 1437 AATCTGCTGATTCATATGCTGTTAAATGGCTTGAATCTGAGAAATTTGAGGTTTCTC 1496
 DB 375 AATAAGTTGTTTTTTTATTTTATATGTTTATGTTTAAAGTTGTTTATGTTATTTT 434
 QY 1497 AAGTAATTTGATCATCAGAACATGATGATCTGATTTCTCAAGTGAATTTTAC 1556
 DB 435 TTGATGATATGAGTT---ATTATATATATATGTTTATATATATATATATATAT 490
 QY 1557 GGTTCCTTCTTAAATTTAGATATATATATATATATATATATATATATATATAT 1616
 DB 491 AATATTTTATATATGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 550
 QY 1617 GATPACTGTCGCAAGTCAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1676
 DB 551 GTAATTTTGTAAATTTGTTAGATTTGTTTATATATATATATATATATATATAT 610
 QY 1677 TTGATGTTATGTTGTTGTTGTTGTT 1703
 DB 611 TTTTATATATTTTATATGTTTATTTT 637

RESULT 14

ABL32357
 ID ABL32357 standard; DNA; 9642 BP.

AC ABL32357;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 330.

XX Human; immune system disease; cytosine methylation; antiaesthetic;
 XX antiarteriosclerotic; antianemic; cytosolic; nootropic;
 XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 XX antirheumatic; antiarthritic; antidiabetic; antiparasitic;
 XX antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
 XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 XX gene; ds.

OS Homo sapiens.

PN WO200200928-A2.

PD 03-JAN-2002.

XX

PF 02-JUL-2001; 2001WO-BP07537.

XX 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation.

PS Claim 1; SEQ ID NO 330; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/bulcerative bowel
 CC diseases. The present sequence is a gene of the invention.

SO Sequence 9642 BP; 2381 A; 88 C; 2349 G; 4824 T; 0 other;

Query Match 2.3%; Score 62.6; DB 24; Length 9642;
 Best Local Similarity 49.5%; Pred. No. 0.00018;
 Matches 161; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 1386 TTTTTCAGATCTGAGCTGTTTATATGATGTTTATGATGTTGATGTTGATGTTG 1445
 DB 3491 TTTTTCAGATCTGAGCTGTTTATATGATGTTTATGATGTTTATGATGTTGATG 3550
 QY 1446 ATTCAATGCTGTTAAATGAGTTGAATCTGAGAAATTTGAGGTTTCTCAAGTAAT 1505
 DB 3551 TTTTATTTTATTTTAAATGTTTAAAGATGTTTATATGTTGATTTTATTTATTTGA 3550
 QY 1506 TTGAATCATCAGAACTATGATGATCTGATTTCTCAAGTGAATTTAGGTTTCTT 1565
 DB 3611 TAAATTTATTAAGATATATTTGTTATATTTCTGATTTAAATTTAAATATATTT 3670
 QY 1566 TCTAATTTAGAGTTATTTATGATGCTAAAGCTTAAATCTTTATATGATGATCTGG 1625
 DB 3671 TTATATTAAGTGGTTTATTTTATATTTATTTTATTTTATTTTATTTTATTTTGA 3720
 QY 1626 TCCAAAGTATGATGATGCTGTTCTTTGCTTACTGATGATGATGATGATGATG 1685
 DB 3731 TTTTATTTTATTTGTTTAAAGTATGATATATTTATTTATTTATTTATTTTGA 3790
 QY 1686 TATGTTTGTGTTTGTGAGTAT 1710
 DB 3791 TTTTATTTTATTTTATTTTGTGAATGT 3815

RESULT 15

AB210194/c
 ID AB210194 standard; DNA; 6343 BP.

AC AB210194;

DT 16-JAN-2003 (first entry)

DE Haematopoietic cell proliferation disorder related DNA sequence #334.

XX Human; haematopoietic cell proliferation disorder; cytostatic;
 XX gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 XX cytosine methylation state; gene; ds.

OS Homo sapiens.

XX

PN WO20027272-A2.
 XX 03-OCT-2002.
 PD 26-MAR-2002, 2002WO-EP03401.
 XX 26-MAR-2001; 2001US-278333P.
 PR (EPIC-) EPIGENOMICS AG.
 XX Berlin K, Braun A, Distler J, Guecig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu B;
 PI Lewin A, Lipsecher E, Mater S, Model F, Mueller V, Otto T;
 PI Pellet C, Schwöpe I, Ziebarth H;
 DR MPI; 2003-018942/01.
 XX Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent
 PT that distinguishes between methylated and non-methylated CpG
 PT dinucleotides -
 XX Claim 28; SEQ ID 334; 117pp; English.
 PS
 CC The present invention describes a method for detecting and
 CC differentiating between hematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. AB209861 to AB21118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used for
 CC differentiating between healthy hematopoietic cells and proliferative
 CC disorder hematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of hematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of hematopoietic cell proliferation disorder related
 CC DNA sequences. The nucleotide sequences from the present invention can
 CC also be used for detecting a predisposition to differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC hematopoietic cell proliferative disorders. The present method enables
 CC a highly specific classification of hematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients.
 XX
 XX Sequence 6343 BP; 1685 A; 0 C; 1623 G; 3035 T; 0 other;
 SQ
 Query Match 2.3%; Score 61.6; DB 25; Length 6343;
 Best Local Similarity 43.1%; Pred. No. 0.00026;
 Matches 295; Conservative 0; Mismatches 389; Indels 0; Gaps 0;
 QY 89 AGACATTAGTACGCTCCGACCTTCTACAAATAAAGCTCATATCCAAAATATT 148
 DB 3223 ACACCTTAATTAATACAAAATAATATTTCCAACTAAAAATAAAAAATTTTCAA 3164
 QY 149 ACATGATCATATCATATATATCCCGCCGAACTTGTATGGACAAACCTGTAACCCCT 208
 DB 3163 TTCTAAACCTCCAAAATAATATCTCTTTTCAATTAAACAACCTTACTTCTTAACAA 3104
 QY 209 TTTTCCTTTTATGTTCAATGACTATACAGTTTGGTTATGATACATATAATGATG 268
 DB 3103 TTTAAAAATCAACATATCTATAATAATATTTTCTTACTCTTTATATAATATTT 3044
 QY 269 GACCCAGCAATTATCCAAATTTGGATATAGATACATAAGCTTAAATCAACATGTAA 328
 DB 3043 ATCTTAACCTAAACCAAAAATACTTTACATATATAATATAATCAAAAATTCAT 2984
 QY 329 CCAAACTAAATCTTTATAGACATAGTAAGTATTCACCAATCTTTATATCATTTGT 388
 DB 2983 ACATTTCTAAAAATATTTTATAAAAATAATTAATTAATAAAAAATTTCTTACATTTAT 2924

QY 389 AAGTACGAAGAGGTAAAAAAGAGAGCCAGCTGATCATACAACTAATCAGACAAA 448
 DB 2923 TTTAATCTATCATATCAAACTAAAAATATCTTTATATCAACAAAATATCATTTTAAACAAA 2864
 QY 449 AGTAGTCAAAATGTTTCTTAAAGTGAGTTTGTATGCAAGAAAAGTGATTAATTTTAA 508
 DB 2863 ATCCATCTAAATTCAAAATACCAATATATCAAAAACAAAACATTAACAAAATTA 2804
 QY 509 TTGAATATATCATATGATGATGATATATCAACAGGCTTCAGTATATAAATTTTAA 568
 DB 2803 TTAATAATTTTATTTAT 2744
 QY 569 ACCAATCAAAATGTTGTTTTCGTCACAAAGTAAGGACCACTGAGAAAGCGT 628
 DB 2743 AAAATTTAATTAACCACTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2684
 QY 629 CCCAGTGTCTCCCTCTTCTTTCTCTGTATTTTGGTCATTTACTATTTACATAC 688
 DB 2683 TAAATATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2624
 QY 689 TCACAGAAAAAAGATTAAGACATTAACACAGTTACTAAGCGTATATCTCTG 748
 DB 2623 AACCTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2564
 QY 749 CACCTTACATACACCTCTTATAT 772
 DB 2563 AACATATTTTAAATTTCTTAAAT 2540

Search completed: December 19, 2003, 20:57:01
 Job time : 687 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 20:30:29 ; Search time 5408 Seconds

(Without alignments)
12075.837 Million cell updates/sec

Title: US-09-995-938A-2

Perfect score: 2687
Sequence: 1 accttgccttcctcaatc.....atcttcgtagataaaaaa 2687

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 segs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0*

Maximum Match 100*
Listing first 45 summaries

Database :

EST:*

| | |
|-----|----------|
| 1: | em_ests: |
| 2: | em_ests: |
| 3: | em_ests: |
| 4: | em_ests: |
| 5: | em_ests: |
| 6: | em_ests: |
| 7: | em_ests: |
| 8: | em_ests: |
| 9: | em_ests: |
| 10: | em_ests: |
| 11: | em_ests: |
| 12: | em_ests: |
| 13: | em_ests: |
| 14: | em_ests: |
| 15: | em_ests: |
| 16: | em_ests: |
| 17: | em_ests: |
| 18: | em_ests: |
| 19: | em_ests: |
| 20: | em_ests: |
| 21: | em_ests: |
| 22: | em_ests: |
| 23: | em_ests: |
| 24: | em_ests: |
| 25: | em_ests: |
| 26: | em_ests: |
| 27: | em_ests: |
| 28: | em_ests: |
| 29: | em_ests: |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 717.6 | 26.7 | 941 | B12579 | B12579 F22H5-T7 IG |
| 2 | 641.4 | 23.9 | 685 | AV782667 | AV782667 AV782667 |
| 3 | 473 | 17.6 | 639 | AV822043 | AV822043 AV822043 |
| 4 | 462 | 17.2 | 463 | AV440785 | AV440785 AV440785 |

| | | | | | | |
|----|-------|------|------|----|-----------|---------------------|
| 5 | 442 | 16.4 | 539 | 9 | AV442375 | AV442375 AV442375 |
| 6 | 416.6 | 15.5 | 796 | 13 | B0238047 | B0238047 DB01_1460 |
| 7 | 386.2 | 14.4 | 1619 | 9 | AV440980 | AV440980 AV440980 |
| 8 | 378 | 14.1 | 1648 | 9 | AV454503 | AV454503 AV454503 |
| 9 | 356 | 13.2 | 1585 | 9 | AV551486 | AV551486 AV551486 |
| 10 | 352.8 | 13.1 | 514 | 14 | T42881 | T42881 6144 Lambda |
| 11 | 323.6 | 12.0 | 773 | 28 | BH673471 | BH673471 B0HXG45TR |
| 12 | 313.8 | 11.7 | 568 | 9 | AV782917 | AV782917 AV782917 |
| 13 | 303.8 | 11.3 | 400 | 14 | U74178 | U74178 AVU74178 NA |
| 14 | 288.8 | 10.7 | 307 | 9 | AA389830 | AA389830 OS198 NAC |
| 15 | 280 | 10.4 | 602 | 29 | B2476824 | B2476824 B0NHO1TR |
| 16 | 277.6 | 10.3 | 604 | 10 | BE662806 | BE662806 EST00551 |
| 17 | 277.6 | 10.3 | 772 | 28 | B20877 | B20877 T30M18-T7 T |
| 18 | 270.8 | 10.1 | 537 | 9 | AV1997341 | AV1997341 701553442 |
| 19 | 262.8 | 9.8 | 643 | 14 | CA934625 | CA934625 MTUSCS_P2 |
| 20 | 249.6 | 9.3 | 511 | 28 | BH597004 | BH597004 B0GDP74TF |
| 21 | 227.2 | 8.5 | 427 | 9 | AV542332 | AV542332 AV542332 |
| 22 | 210.4 | 7.8 | 511 | 10 | BF277350 | BF277350 GA_EB002 |
| 23 | 201.8 | 7.5 | 629 | 13 | BQ407744 | BQ407744 GA_EB000 |
| 24 | 201.6 | 7.5 | 541 | 12 | B116769 | B116769 F073P54Y |
| 25 | 200.4 | 7.5 | 692 | 28 | BH500787 | BH500787 B0GMO41TR |
| 26 | 196 | 7.3 | 456 | 29 | CNS00ND | AL082567 Arabidops |
| 27 | 195.8 | 7.3 | 582 | 14 | CA933175 | CA933175 MTUSCS_P4 |
| 28 | 194.6 | 7.2 | 590 | 9 | AV822219 | AV822219 AV822219 |
| 29 | 184.4 | 7.2 | 801 | 12 | B1433025 | B1433025 EST535786 |
| 30 | 183.8 | 7.2 | 812 | 14 | CB292897 | CB292897 UCRC501_0 |
| 31 | 183.4 | 7.2 | 315 | 28 | BH634176 | BH634176 SALK_0448 |
| 32 | 183.2 | 7.2 | 599 | 10 | BG447420 | BG447420 GA_EB004 |
| 33 | 191 | 7.1 | 359 | 9 | AV553607 | AV553607 AV553607 |
| 34 | 190.8 | 7.1 | 626 | 13 | BQ407763 | BQ407763 GA_EB000 |
| 35 | 188.2 | 7.0 | 688 | 28 | BH533355 | BH533355 B0GMO62TF |
| 36 | 186.8 | 7.0 | 504 | 12 | B1137968 | B1137968 F096P47Y |
| 37 | 185.2 | 6.9 | 451 | 9 | AV532273 | AV532273 AV532273 |
| 38 | 183 | 6.8 | 359 | 28 | BH645352 | BH645352 B0XHU29TF |
| 39 | 181.8 | 6.8 | 766 | 29 | B2476333 | B2476333 B0NBU61TF |
| 40 | 178.4 | 6.6 | 661 | 28 | BH428064 | BH428064 B0GMO55TR |
| 41 | 173.8 | 6.5 | 458 | 12 | B116953 | B116953 F077P40Y |
| 42 | 173.8 | 6.5 | 531 | 9 | AV734606 | AV734606 SK97A09_Y |
| 43 | 172.4 | 6.4 | 400 | 9 | AV807059 | AV807059 AV807059 |
| 44 | 170.8 | 6.4 | 481 | 9 | AV1992723 | AV1992723 701493393 |
| 45 | 166.8 | 6.2 | 606 | 13 | BQ591032 | BQ591032 E012599-0 |

ALIGNMENTS

RESULT 1

LOCUS B12579

DEFINITION F22H5-T7 IGF Arabidopsis thaliana genomic clone F22H5, genomic survey sequence.

ACCESSION B12579

VERSION B12579.1 GI:2093659

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 941)

AUTHORS Feng J., Dewar K., Buehler E., Kim C., Li Y., Shinn P., Sun H. and Ecker J.

TITLE BAC End Sequences at ATGC

JOURNAL Unpublished

COMMENT Other_GSSs: F22H5-SP6

Contact: Ecker J.

Arabidopsis Thaliana Genome Center

University of Pennsylvania

Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104

Tel: 215-898-9384

Fax: 215-898-8780

Email: jecker@atgenome.bio.upenn.edu

Seq primer: T7
 Class: BAC ends
 High quality sequence start: 94
 High quality sequence stop: 469.
 Location/Qualifiers
 1..941
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="F2H5"
 /sex="hermaphrodite"
 /clone_lib="IGF"
 /note="Vector: BelobacII; Site_1: EcoRI; Site_2: EcoRI;
 Produced by Thomas Altmann"

BASE COUNT 230 a 213 c 157 g 324 t 17 others

Query Match 26.7%; Score 717.6; DB 28; Length 941;
 Best Local Similarity 90.1%; Pred. No. 3e-104;
 Matches 799; Conservative 0; Mismatches 81; Indels 7; Gaps 3;

1448 TTCAATTGGTTAAATGGGTTGAATCTGAGAAATTTGAGGTTTCTCAAGTGAATT 1507
 49 TTCAATTGGTGGCCCGGGTTGAATCTGAGAAATTTGAGGTTTCTCAAGTGAATT 108
 1508 GAATATCGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1567
 109 GAATATCGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 168
 1568 TAATTTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1627
 169 TAATTTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 228
 1628 CAAAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1687
 229 CAAAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 288
 1688 TTGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1747
 289 TTGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 348
 1748 CTTCACTCGAGTAACTCCATATTCATCAAGAACAGAGCCCTCTTATCAGCCTTTC 1807
 349 CTTCACTCGAGTAACTCCATATTCATCAAGAACAGAGCCCTCTTATCAGCCTTTC 408
 1808 AAAGTCCCATCCATCTTAACGAAGTAAAGCCGCTCTTCTTATCAGCCTTTC 1867
 409 AAAGTCCCATCCATCTTAACGAAGTAAAGCCGCTCTTCTTATCAGCCTTTC 468
 1868 GCGGGAACCAATTAACAACATGCTCTCAATTTCTCCCTTTCTCAGAAATGGTGGCA 1927
 469 GCGGGAACCAATTAACAACATGCTCTCAATTTCTCCCTTTCTCAGAAATGGTGGCA 528
 1928 TTCTCTTCT 1987
 529 TTCTCTTCT 588
 1988 CATGCGGCACTTTAAGAACCCGAAACCGTTGCTTAATCTGGAATCTATCGCTAAGCAT 2047
 589 CATGCGGCACTTTAAGAACCCGAAACCGTTGCTTAATCTGGAATCTATCGCTAAGCAT 648
 2048 CCAATGGCACTTCTAAGAACCCGAAACCGTTGCTTAATCTGGAATCTATCGCTAAGCAT 2107
 649 CCAATGGCACTTCTAAGAACCCGAAACCGTTGCTTAATCTGGAATCTATCGCTAAGCAT 708
 2108 CCAATGGCACTTCTAAGAACCCGAAACCGTTGCTTAATCTGGAATCTATCGCTAAGCAT 2167
 709 CCAATGGCACTTCTAAGAACCCGAAACCGTTGCTTAATCTGGAATCTATCGCTAAGCAT 768
 2168 ATGAGTCTGACTTCTCACTG- TTGATTTGCTGCTAATGG- ATGAGTCTGAGAGTTTG 2224
 769 ATGAGTCTGACTTCTCACTGTTGATTTGCTGCTAATGGGATTAATCTTCAAAAGTTTG 828

2225 CACA-----ACACAGCATTCTCTGCTCTGATAGTGGCAACCTCTCTCACTTCAATCTT 2280
 829 CACANNACAGCAATTTCTCTGCTCTGATAGTGGCAACCTCTCTCACTTCAATCTT 888
 2281 GTGAACCTGCGGCTCAGCAGATGCTCCAAATGAGTGGCTTCA 2327
 889 GTTGAACCTGCGGCTCAGCAGATGCTCCAAATGAGTGGCTTCA 935

RESULT 2
 AV782667/c 685 bp mRNA linear EST 28-MAR-2002
 LOCUS AV782667 RAF14 Arabidopsis thaliana cDNA clone RAF104-20-E20 3'
 DEFINITION mRNA sequence.
 AV782667
 AV782667.1 GI:19801457
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; euroside II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 685)
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Saitou, M., Nakajima, M.,
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.,
 and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
 Unpublished
 CONTACT: Motoki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: msek@rtc.riken.go.jp

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). This clone is in a
 modified Bluescript vector as a set/XhoI insert. Please visit our
 web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for
 further details.

FEATURES
 source
 location/Qualifiers
 1..685
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF104-20-E20"
 /dev_stage="rosette plants"
 /lab_host="SOUR"
 /clone_lib="RAF14"
 /note="Site 1: SclI; Site 2: XhoI; subjected to
 cold-treated(1,2,5,10,24 hr)"

BASE COUNT 217 a 134 c 147 g 187 t
 ORIGIN
 Query Match 23.9%; Score 641.4; DB 9; Length 685;
 Best Local Similarity 99.7%; Pred. No. 4.1e-92;
 Matches 653; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

2034 TATGCTAAGAAATCCATGCGCATTTGCTTAACAAATCAATGAGGCTTTTAATATTCCTT 2093
 685 TATGCTAAGAAATCCATGCGCATTTGCTTAACAAATCAATGAGGCTTTTAATATTCCTT 626
 2094 CTATGCGGTTCTGCACTCTGATGCGACATGCGCACAGGTTTCAATCCCGGCTAC 2153
 625 CTATGCGGTTCTGCACTCTGATGCGACATGCGCACAGGTTTCAATCCCGGCTAC 566
 2154 TATACCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2213
 565 TATACCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 506
 2214 TCAGAAATTTGCAACAACAGCCATTTCTGCTCTATGATGCGCAACCTCTCTCACTT 2273

Db 505 TCAGAGTTGCAACAACAGCCATTCTGCTCATGTTGGTCCAACTCTCTACTT 446
QY 2274 CAATCTTGGAACCTGCGCTCAGCAGATGTCCTCAAAATACGTCCTTCCAGAGAT 2333
Db 445 CAATCTTGGAACCTGCGCTCAGCAGATGTCCTCAAAATACGTCCTTCCAGAGAT 386
QY 2334 TGGTCAAGCTCGAGTTTAAATTTGAAATAGCAAGTTAAACCTCGGAGAGAGAG 2393
Db 385 TGGTCAAGCTCGAGTTTAAATTTGAAATAGCAAGTTAAACCTCGGAGAGAGAG 326
QY 2394 GATACATGATGTTGGATGAGAGATCTTGAGCTTACCTGGAATGGAAGCTCGTGG 2453
Db 325 GATACATGATGTTGGATGAGAGATCTTGAGCTTACCTGGAATGGAAGCTCGTGG 266
QY 2454 TTGACATTAACAACATAGGCAAAACCAATGCGATGTCATGGAATATGAGAACTAATCC 2513
Db 265 TTGACATTAACAACATAGGCAAAACCAATGCGATGTCATGGAATATGAGAACTAATCC 206
QY 2514 TCTTGAGTATTTCTTCTTGCTGCTGAGATTTGGAATCTTATGGAATCTCATATGTTCTT 2573
Db 205 TCTTGAGTATTTCTTCTTGCTGCTGAGATTTGGAATCTTATGGAATCTCATATGTTCTT 146
QY 2574 CACTTATTAACAACATAGGCAAAACCAATGCGATGTCATGGAATATGAGAACTAATCC 2632
Db 145 CACTTATTAACAACATAGGCAAAACCAATGCGATGTCATGGAATATGAGAACTAATCC 86
QY 2633 GTGGTATTAACAACATAGGCAAAACCAATGCGATGTCATGGAATATGAGAACTAATCC 2687
Db 85 GTGGTATTAACAACATAGGCAAAACCAATGCGATGTCATGGAATATGAGAACTAATCC 31

RESULT 3
AV822043 639 bp mRNA linear EST 01-APR-2002
LOCUS AV822043 RAF14 Arabidopsis thaliana cDNA clone RAF14-20-E20 5'
DEFINITION mRNA sequence.
ACCESSION AV822043
VERSION AV822043.1 GI:19864075
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eustoids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 639)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Saitou,M., Nakajima,M., Arai,K., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arai,K., T., Shibata,K., Shingawa,A., Muramatsu,M., Hayashizaki,Y., and Shinozaki,K.
COMMENT Large scale analysis of Arabidopsis full-length cDNA (2002b)
UNPUBLISHED
CONTACT: Motoaki Seki
PLANT Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: meeki@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SecI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.
FEATURES
Source Location/Qualifiers
1..639
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF14-20-E20"
/db_stg="roseette plants"
/lab_host="SOLR"
/clone_id="RAF14"

BASE COUNT 168 a 144 c 153 g 174 t
ORIGIN
Query Match 17.6%; Score 473; DB 9; Length 639;
Best Local Similarity 100.0%; Pred. No. 2.4e-65;
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 898 AGATGAAAAAATATATATTTGTTGAGAGAAGAGAGATTTCTTCTTGATTC 957
Db 2 AGATGAAAAAATATATATTTGTTGAGAGAAGAGAGATTTCTTCTTGATTC 61
QY 958 AGGGAAGAAAAACGATTTCTGAGACCTAATCTTCACTCTCTTCTTCTT 1017
Db 62 AGGGAAGAAAAACGATTTCTGAGACCTAATCTTCACTCTCTTCTTCTT 121
QY 1018 CATCAGCTACGTTTCAACAATCTTCCACCCTATTCAGAGCTCTCCGGAATTC 1077
Db 122 CATCAGCTACGTTTCAACAATCTTCCACCCTATTCAGAGCTCTCTCCGGAATTC 181
QY 1078 GAGGGGTGGTTGTTGTTTCCGATGACTTGGATGAGCTACGTCAGATCAGAGC 1137
Db 182 GAGGGGTGGTTGTTGTTTCCGATGACTTGGATGAGCTACGTCAGATCAGAGC 241
QY 1138 TGCAGCTGCGGCGCGGACAGCGGCGAGAGAGACCGTCTGGAGAGAGAGAA 1197
Db 242 TGCAGCTGCGGCGCGGACAGCGGCGAGAGAGACCGTCTGGAGAGAGAGAA 301
QY 1198 TAATCGAGAGAGAAAGACGAGAGAGAGCTGATGCGAGATATACACTGGGCTTAA 1257
Db 302 TAATCGAGAGAGAAAGACGAGAGAGAGCTGATGCGAGATATACACTGGGCTTAA 361
QY 1258 AGCTCAAGGTATTAATTTGCTTAAACATTTGATTAATGAACTCTTAAAGCTCT 1317
Db 362 AGCTCAAGGTATTAATTTGCTTAAACATTTGATTAATGAACTCTTAAAGCTCT 421
QY 1318 TTGTGTTGAAGCTGTTGGTTGTTGAAGAAGATGATCTTATGCAAGG 1370
Db 422 TTGTGTTGAAGCTGTTGGTTGTTGAAGAAGATGATCTTATGCAAGG 474

RESULT 4
AV440785/c 463 bp mRNA linear EST 14-NOV-2000
LOCUS AV440785 Arabidopsis thaliana above-ground organ two to six-week
DEFINITION old Arabidopsis thaliana cDNA clone Ap209d10_E 3', mRNA sequence.
ACCESSION AV440785
VERSION AV440785.1 GI:7611157
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eustoids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 463)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S., and Tabata,S.
COMMENT A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
DNA Res. 7, 175-180 (2000).
JOURNAL MEDLINE 20363093
PUBMED 10907847
CONTACT: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
Source Location/Qualifiers
1..463
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/db_stg="Columbia"

```
/db_xref="taxon:3702"
/clone="AP209b10_f"
/issue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_1b="Arabidopsis thaliana above-ground organ two to
six-week old"
/notes="Vector: pBluescriptII SK-, Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      139 a      96 c      92 g      135 t      1 others
ORIGIN
Query Match      17.2%; Score 462; DB 9; Length 463;
Best Local Similarity 99.8%; Pred. No. 1.4e-63;
Matches 462; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2224 GCACAACACAGCCATCTCTGCTATGTCGCCAAGCTCTCTCACTTCAATCTTGTG 2283
DB 463 GCACAACACAGCCATCTCTGCTATGTCGCCAAGCTCTCTCACTTCAATCTTGTG 404
QY 2284 AAACCTGCCCTCAGAGATGTCGCAATATCTGCTGCTTCCAGAGATGGTCAAGC 2343
DB 403 AAACCTGCCCTCAGAGATGTCGCAATATCTGCTGCTTCCAGAGATGGTCAAGC 344
QY 2344 TCTGAGTTAAATTTGAGATAGCCAAAGTTAAACCTGGAGAGAGAGATGATGAT 2403
DB 343 TCTGAGTTAAATTTGAGATAGCCAAAGTTAAACCTGGAGAGAGAGATGATGAT 284
QY 2404 GTGGGTATGAGAGATCTTACCTTACCTTGAAGAGAGAGAGATGATGATGAT 2463
DB 283 GTGGGTATGAGAGATCTTACCTTACCTTGAAGAGAGAGAGATGATGATGAT 224
QY 2464 CAACCTAGGCAACCCCAATGAGATGATGATGATGATGATGATGATGATGAT 2523
DB 223 CAACCTAGGCAACCCCAATGAGATGATGATGATGATGATGATGATGATGAT 164
QY 2524 TTTCTTCTCTGTCAGAGATTTGGATCTTTATGAGATCTCAATGCTTCTTAT 2583
DB 163 TTTCTTCTCTGTCAGAGATTTGGATCTTTATGAGATCTCAATGCTTCTTAT 104
QY 2584 CCAATATGCTGCCCAAGCCTTCTCCATGAGAGATGAGATGAGATGAGATGAT 2643
DB 103 CCAATATGCTGCCCAAGCCTTCTCCATGAGAGATGAGATGAGATGAGATGAT 44
QY 2644 ATTCAATGATTTGGTTCAAAAGCATTTATTGTAGATATAAAAA 2686
DB 43 ATTCAATGATTTGGTTCAAAAGCATTTATTGTAGATATAAAAA 1

RESULT 5
AV442375 539 bp mRNA linear EST 14-NOV-2000
LOCUS AV442375 Arabidopsis thaliana above-ground organ two to six-week
DEFINITION old Arabidopsis thaliana cDNA clone AP209b10_r 5', mRNA sequence.
ACCESSION AV442375
VERSION AV442375.1 GI:7612783
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 539)
REFERENCE Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
AUTHORS A large scale analysis of cDNA in Arabidopsis thaliana: Generation
TITLE of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 252-0812, Japan
```

```
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..539
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="AP209b10_f"
/issue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_1b="Arabidopsis thaliana above-ground organ two to
six-week old"
/notes="Vector: pBluescriptII SK-, Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      141 a      118 c      137 g      143 t
ORIGIN
Query Match      16.4%; Score 442; DB 9; Length 539;
Best Local Similarity 100.0%; Pred. No. 2.1e-60;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 929 AGAAGAGAGATTTCTTCTTCTTCATCCAGGAGAGAAAGGATTCCTCGAGAC 988
DB 1 AGAAGAGAGATTTCTTCTTCTTCATCCAGGAGAGAAAGGATTCCTCGAGAC 60
QY 989 TAACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1048
DB 61 TAACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 1049 ACCTATTCAGACCTCTCCGAGATTTGAGAGATTTGAGAGATTTGAGAGAT 1108
DB 121 ACCTATTCAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 1109 TCGATGAGATACCTTCGACATCAGACGTCGACGTCGAGCGGCGAGCGAGG 1168
DB 181 TCGATGAGATACCTTCGACATCAGACGTCGACGTCGAGCGGCGAGCGAGG 240
QY 1169 AGAAGCCGCTGTCGAGAGAGAGAGAGATTAATGAGAGAGAGAGAGAGAGCT 1228
DB 241 AGAAGCCGCTGTCGAGAGAGAGAGAGATTAATGAGAGAGAGAGAGAGAGCT 300
QY 1229 GTAGCTGCAATATATACCTGGCTTGAAGCTGAAGCTGAATTAATTTGCTTAACAT 1288
DB 301 GTAGCTGCAATATATACCTGGCTTGAAGCTGAAGCTGAATTAATTTGCTTAACAT 360
QY 1289 TGTGATTAATGAGAGATCCTTAAGCTCTTGTGTGAGAGCTGCTGTTGAAGAA 1348
DB 361 TGTGATTAATGAGAGATCCTTAAGCTCTTGTGTGAGAGCTGCTGTTGAAGAA 420
QY 1349 GATGTAATTAATTCGCAAG 1370
DB 421 GATGTAATTAATTCGCAAG 442

RESULT 6
BU238047 796 bp mRNA linear EST 06-SEP-2002
LOCUS BU238047/c
DEFINITION Ds01.14f06 A/Ds01.14f06 ECORC cold stressed Flaxseed seedlings
DEFINITION Descurainia sophia cDNA clone Ds01.14f06, mRNA sequence.
ACCESSION BU238047
VERSION BU238047.1 GI:22749872
KEYWORDS EST.
SOURCE Descurainia sophia
ORGANISM Descurainia sophia
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Descurainia.
1 (bases 1 to 796)
REFERENCE Singh, V.A., Piche, C., Couroux, P., De Moors, A., Harris, L.J., Hattori,
AUTHORS J.I., Ouellet, T., Robert, L.S., Sproule, D. and Tinker, N.A.
TITLE Expressed Sequence Tags from Cold-Stressed Descurainia sophia
JOURNAL Seedlings
Unpublished
```

COMMENT

Contact: Singh, J.A.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-food Canada
 3610 Hwy 1, Central Experimental Farm, Ottawa, Ontario, K1A
 0C6, Canada
 Tel: (613) 759-1662
 Fax: (613) 759-1701
 Email: singhja@agr.ca

FEATURES

Source

Location/Qualifiers
 1. 796
 /organism="Descurainia sophia"
 /mol_type="mRNA"
 /db_xref="taxon:89411"
 /clone="D801.14f06"
 /cissue_type="leaf, stem"
 /dev_stage="1-month seedlings - 1 cm tall - 8 leaf"
 /clone_1lb="D801_AmFC_ECORC_cold_stressed_Plixweed_seedlin
 gs"
 /note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: Eco RI;
 Site 2: Xho I. Plants were grown for 1 month at 200C/16
 hrs light/day (average 8 leaves, 1 cm tall, weight
 0.02g/plant). Then they were exposed to 20C, 12 hrs
 light/day, for 1 week. Library prepared by C. Piche using
 Stratagene kit."

BASE COUNT 221 a 150 c 211 g 207 t 7 others
 ORIGIN

Query Match 15.5%; Score 416.6; DB 13; Length 796;
 Best Local Similarity 73.8%; Pred. No. 2.1e-56;
 Matches 553; Conservative 3; Mismatches 187; Indels 6; Gaps 2;

1831 GTTAGCCGCTCTTTCATCATTTCCGAGTCTTCTGCGGTGAACCAATTAACAATC 1830
 796 GTAGTCCATGCTTCATCATTTCCGAGTCTTCTGCGGTGAACCAATTAACAATC 737
 1891 TCCCTCATCTCTTCT 1950
 736 TCC---ACATCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 680
 1951 AGAATCTCAACAGTTGTCAGTACCCGAGTCTCTATGCGGACTTCTAGAACCCG 2010
 679 AGATCTCAACAGTTGTCAGTACCCGAGTCTCTATGCGGACTTCTAGAACCCG 620
 2011 AAACCTGCTCTACTGGAATCTATGCTATGCAATCTATGCGGACTTCTAGAAC 2067
 619 AAACCTGCTCTACTGGAATCTATGCTATGCAATCTATGCGGACTTCTAGAAC 560
 2068 TCAATGCGCTCTTAAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2127
 559 TCAATGCGCTCTTAAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 500
 2128 CGCAGCAGTTTCTATACCTGCTACTATACCTGATGATGATGATGATGATGATG 2187
 499 CATGCGAGTTTCTATACCTGCTACTATACCTGATGATGATGATGATGATGATG 440
 2188 GTTGATCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 2247
 439 GTTGATCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 380
 2248 TCAATGCTGCT 2307
 379 TCAATGCTGCT 320
 2308 CCAAAATGCTGCT 2367
 319 CCAAAATGCTGCT 260
 2368 CAAATTAACCTGCT 2427
 259 CAAATTAACCTGCT 200
 2428 ACACTTGAATGGAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2487

Db

199 ACACTTGAATGGAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 140

Qy

2488 GTCAATGGAATGGAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2547

Db

139 GAAGATGGAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 80

Qy

2548 ATCTTANGAATCTATGCTCTCAC 2576

Db

79 ATATCTTAGACGAGCTTCTTCTAC 51

RESULT 7

AV440980/C 619 bp mRNA linear EST 14-NOV-2000

LOCUS AV440980 Arabidopsis thaliana above-ground organ two to six-week
 DEFINITION old Arabidopsis thaliana cDNA clone AP218b09_f3', mRNA sequence.

ACCESSION

AV440980

VERSION

AV440980.1 GI:7611357

KEYWORDS

EST.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana (thale cress)

REFERENCE

Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 DNA Res. 7, 175-180 (2000)

AUTHORS

Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.

TITLE

A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries

JOURNAL

DNA Res. 7, 175-180 (2000)

MEDLINE

20363093

PUBMED

10907847

COMMENT

Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

Location/Qualifiers
 1. 619
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="AP218b09_f3"
 /cissue_type="aboveground organs"
 /dev_stage="two to six-week old"
 /clone_1lb="Arabidopsis thaliana above-ground organ two to
 six-week old"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"

Source

Location/Qualifiers
 1. 619
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="AP218b09_f3"
 /cissue_type="aboveground organs"
 /dev_stage="two to six-week old"
 /clone_1lb="Arabidopsis thaliana above-ground organ two to
 six-week old"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 176 a 129 c 148 g 166 t
 ORIGIN

Query Match 14.4%; Score 386.2; DB 9; Length 619;
 Best Local Similarity 83.9%; Pred. No. 1.5e-51;
 Matches 449; Conservative 0; Mismatches 83; Indels 3; Gaps 1;

1959 AAACAGTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2018
 619 AAACAGTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 560
 2019 GCTTAATGGAATCTTATGCAATGATGATGATGATGATGATGATGATGATG 2075
 559 GCTTAATGGAATCTTATGCAATGATGATGATGATGATGATGATGATGATG 500
 2076 GCTTAATGGAATCTTATGCAATGATGATGATGATGATGATGATGATGATG 2135
 499 TTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 440
 2136 GTTTCATACCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATG 2195
 439 GTTTCATACCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATG 380

2196 TGGTCATTTGATAGCTTTTTCAGAGTTTGCACAAACAGCCATTTCTGCTCTATAGT 2255
Db 379 TGGTCATTTGATAGCTTTTTCAGAGTTTGCACAAACAGCCATTTCTGCTCTATAGT 320
Qy 2256 GCCAACCTTCTCTACCTTCAATCTTGTGAACCTGCGCTCAGCAGATGTCTCCAAATAC 2315
Db 319 GCCAACCTTCTCTACCTTCAATCTTGTGAACCTGCGCTCAGCAGATGTCTCCAAACAC 260
Qy 2316 TGGTCCTTCCAGAGATTTGGTCAAGCTGTAGTTTAAATTTGAGAAATGCCAAGTTAA 2375
Db 259 AGCAGCAATTCAGAGATTTGGTCAAGCTGTAGTTTAAATTTGAGAAATGCCAAGTTAA 200
Qy 2376 ACCCTGGGAAGAGAGAGATATGATGTGGTATGAGAGATCTTGTAGCTTACCTGG 2435
Db 199 GCCATGGGAAGAGAGAGATATGATGTGGTATGAGAGATCTTGTAGCTTACCTGG 140
Qy 2436 AATGGGAAGCTGTGTGTGATGATTAACAACCTAGCCAAACCCAAATGGCATGTC 2490
Db 139 AATGTGAAGCTGTGTGTGATGATTAACAACCTAGCCAAACCCAAATGGCATGTC 85

RESULT 8
AV545003/c 648 bp mRNA linear EST 07-SEP-2000
LOCUS AV545003 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
DEFINITION CDNA clone R267C11F 3', mRNA sequence.
ACCESSION AV545003
VERSION AV545003.1 GI:8716417
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 648)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
LOCATION/Qualifiers
FEATURES
source
1..648
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="R267C11F"
/tissue_type="roots"
/clone_lib="Arabidopsis thaliana roots Columbia"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 187 a 136 c 151 g 174 t
ORIGIN
Query Match 14.1%; Score 378; DB 9; Length 648;
Best Local Similarity 84.1%; Pred. No. 3e-50;
Matches 439; Conservative 0; Mismatches 80; Indels 3; Gaps 1;

Qy 1972 GTTACCCCAACCGGTCTCATGCGGACTTCTAAGAACCCGAAACCGTGGCTTAAGTGGAA 2031
Db 648 GTTACCCCAACCGGTCTCATGCGGACTTCTAAGAACCCGAAACCGTGGCTTAAGTGGAA 589

Qy 2032 TTTATGCTTAAGCAATCATGGCCAT--TGCTAAACAATCAATGGCGCTTTTATTTAT 2088
Db 588 TCTTTTACCAACAATCATGCTCATGCTGCTAAACAGTCAATGATCTTTTGAACATAC 529

2089 CTTTCTATGCGGTTTCTGCACTGCTGATGCCACATGCGCACCAAGTTTCATACCTTG 2148
Db 528 CCGTTTATGCGGTTTCTGCACTGCTGATGCCACATGCGCACCAAGTTTCATACCTTG 469
Qy 2149 GCTACTATACCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2208
Db 468 GCTACTATACCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 409
Qy 2209 AGCTTCAGAGTTGGTGCACAAACAGCCATCTGCTCTATAGTGGCCAACTCTCT 2268
Db 408 AGCTTCAGAGTTGGTGCACAAACAGCCATCTGCTCTATAGTGGCCAACTCTCT 349
Qy 2269 ACCTTCATCTGTGAACCTGCGCTCAGCAGATGTCTTCAATATCTGCTGCTTCAA 2328
Db 348 ACCTTCATCTGTGAACCTGCGCTCAGCAGATGTCTTCAATATCTGCTGCTTCAA 289
Qy 2329 GAGATTGTCAGAGCTCTGAGTTTAAATTTGAGAAATAGCCAACTTAAACCTGGAGAGA 2388
Db 288 GAGATTGTCAGAGCTCTGAGTTTAAATTTGAGAAATAGCCAACTTAAACCTGGAGAGA 229
Qy 2389 GAGAGATATCATGATGTGGTATGAGAGATCTTGTAGCTTACCTTGAAGAGAGCT 2448
Db 228 GAGAGATATCATGATGTGGTATGAGAGATCTTGTAGCTTACCTTGAAGAGAGCT 169
Qy 2449 CGTGTGACATTAACAATCAGGCAACCCCAATGGCATGTC 2490
Db 168 CATGTTGAGATGAAGTATATCATGAACCTGTATGTCTATGTC 127

RESULT 9
AV551486 585 bp mRNA linear EST 06-SEP-2000
LOCUS AV551486 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
DEFINITION CDNA clone R2127d10R 5', mRNA sequence.
ACCESSION AV551486
VERSION AV551486.1 GI:8722899
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 585)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
LOCATION/Qualifiers
FEATURES
source
1..585
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="R2127d10R"
/tissue_type="roots"
/clone_lib="Arabidopsis thaliana roots Columbia"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 146 a 142 c 138 g 159 t
ORIGIN
Query Match 13.2%; Score 356; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.4e-47;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1015 CTTCATCAGCTACGTCACACATCTTTACCCACCTTATCAAGCTCTCCGGAAGT 1074
 1016 CTTATCAGCTACGTCACACATCTTTACCCACCTTATCAAGCTCTCTCCGGAAGT 60
 1017 TTCCGAGGGGTTGTTGTTGTTTCCCGATGACTTCGATGAGCTACGTCACATCAGC 1134
 1018 TTCCGAGGGGTTGTTGTTGTTTCCCGATGACTTCGATGAGCTACGTCACATCAGC 120
 1019 AGCTCAGCTGCGGCG 1194
 1020 AGCTCAGCTGCGGCG 180
 1021 GAAATATCGGAG 1254
 1022 GAAATATCGGAG 240
 1023 TAGAGCTCAGGCTGATTAATTAATTTGCTTAAACATTTGATTAATTAATTAATTA 1314
 1024 TAGAGCTCAGGCTGATTAATTAATTTGCTTAAACATTTGATTAATTAATTAATTA 300
 1025 TCTTTGTTGAAGCTGTTGGTGTGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1370
 1026 TCTTTGTTGAAGCTGTTGGTGTGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 356
 1027 TCTTTGTTGAAGCTGTTGGTGTGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG
 RESULT 10
 T42881 514 bp mRNA linear EST 07-JAN-1998
 LOCUS 6144 Lambda-PRL2 Arabidopsis thaliana cDNA clone 119H677, mRNA
 DEFINITION
 T42881
 T42881.1 GI:2758126
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 1 (bases 1 to 514)
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
 L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Reetz
 E. and Somerville,C.
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 7846151
 On Jan 7, 1998 this sequence version replaced gi:947520.
 COMMENT
 JOURNAL
 MEDLINE
 PUBMED
 CONTACT: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 2231jrc@lhm.cl.msu.edu
 Seq primer: T7
 FEATURES
 source
 1..514
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="var Columbia"
 /db_xref="taxon:3702"
 /clone="119H677"
 /clone_lib="Lambda-PRL2"
 /note="vector: lambda Zip-Lox, Site_1: Sal, Site_2: Not,
 Lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark-rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques). The vector is BRL's lambda Zip-Lox. The cDNA

BASE COUNT 126 a 133 c 123 g 132 t 20 others
 ORIGIN
 Query Match 13.1%; Score 352.8; DB 14; Length 514;
 Best Local Similarity 92.8%; Pred. No. 3.1e-46; Indels 4; Gaps 3;
 Matches 397; Conservative 0; Mismatches 27;
 929 AGAAGAGAGATTCCTTCTTCTGATTCGAGCGAAGAAAGCGTATTCCTCGTGACAC 988
 1 AGAAGAGAGATTCCTTCTTCTGATTCGAGCGAAGAAAGCGTATTCCTCGTGACAC 60
 989 TAATCTTCACCT 1048
 61 TAATCTTCACCT 120
 1049 ACCTATTCAAAGCT 1108
 121 ACCTATTCAAAGCT 180
 1109 TCGGATGAGCTACGTCGACATCAGCAGCTGCGCGCGCGCGCGCGCGCGCGCGAG 1168
 181 TCGGATGAGCTACGTCGACATCAGCAGCTGCGCGCGCGCGCGCGCGCGCGCGAG 239
 1169 AGAAGCCGTCCTGAGCT 1228
 240 AGAAGCCGTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 299
 1229 GTAGCTCGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1288
 300 GTAGCTCGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 359
 1289 TGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1345
 360 TGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 419
 1346 GAAGATCG 1353
 420 GGAGATTG 427
 RESULT 11.
 BH673471 773 bp DNA linear GSS 19-FEB-2002
 LOCUS BOHXG46TR BO_2_3_KB Brassica oleracea genomic clone BOHXG46,
 DEFINITION
 BOHXG46TR BO_2_3_KB Brassica oleracea genomic clone BOHXG46,
 genomic survey sequence.
 BH673471
 BH673471.1 GI:18740935
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 1 (bases 1 to 773)
 Town,C.D., Van Aken,S., Utechtack,T., Koo,H. and Frazer,C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 unpublished
 JOURNAL
 COMMENT
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-0523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Oaborn.
 Seq primer: TR
 Classes: sheared ends.
 FEATURES
 source
 1..773
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"

AUTHORS Pih,K.T., Park,J.M., Jang,H.J., Kang,S.G., Piao,H.L. and Hwang,I.
TITLE EST of salt inducible mRNA in Arabidopsis thaliana
JOURNAL Unpublished
COMMENT Contact: Kyeong Tae Pih
 Department of Plant Molecular Biology
 Gyeongsang National Univ., Plant Molecular Biology and
 Biotechnology Research Center
 Jinju, Kyungnam 660-701, Korea
 Tel: 82-591-751-5193
 Fax: 82-591-759-9363
 Email: ihwang@ongae.gsnu.ac.kr.

FEATURES
source
 1..400
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="OS158"
 /dev_stage="7 day-old"
 /clone_lib="NaCl-treated Arabidopsis subtraction library"
 /note="Subtracted cDNA library from salt (NaCl)-treated whole plants"

BASE COUNT 96 a 124 c 57 g 121 t 2 others
ORIGIN

Query Match 11.3%; Score 303.8; DB 14; Length 400;
Best Local Similarity 98.8%; Pred. No. 2e-38;
Matches 327; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1792 CTTTATGAGCGTTTCAAGTCCATCCATCTTACCAAGTTGCGCGTCTTTCATCA 1851
 DB 1 CTTTATGAGCGTTTCAAGTCCATCCATCTTACCAAGTTGCGCGTCTTTCATCA 60
 QY 1852 TTCCGAGTCTTCTGCGGTGAACCAATACAGATGCTCTCAATTCCTTC 1911
 DB 61 TTCCGAGTCTTCTGCGGTGAACCAATACAGATGCTCTCAATTCCTTC 120
 QY 1912 CTCAGAAATGGTGGCATTCCTTCTTCTTCCCTCAAGATTCGAAAGTGTCCA 1971
 DB 121 CTCAGAAATGGTGGCATTCCTTCTTCTTCCCTCAAGATTCGAAAGTGTCCA 180
 QY 1972 GTTACCCACCGGTCTCATGCGGATCTTAAGACCGGAACCGTGGCTTAATGCGGA 2031
 DB 181 GTTACCCACCGGTCTCATGCGGATCTTAAGACCGGAACCGTGGCTTAATGCGGA 240
 QY 2032 TCTATGCTAAGCATCATGCGCATTCGTAACATCAATGCGGCTTTAATATTC 2089
 DB 241 TCTATGCTAAGCATCATGCGCATTCGTAACATCAATGCGGCTTTAATATTC 300
 QY 2090 CTTTCTATGCGGTTTCTGACCTGCTAGTCC 2120
 DB 301 CTTTCTATGCGGTTTCTGACCTGCTAGTCC 331

RESULT 14 307 bp mRNA linear EST 23-APR-1997
AA389830
LOCUS OS198 NaCl-treated Arabidopsis subtraction library Arabidopsis
DEFINITION thaliana cDNA 5', mRNA sequence.
ACCESSION AA389830
VERSION AA389830.1 GI:2042816
KEYWORDS EST.
SOURCE Arabidopsis thaliana (chale crese)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryote, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;
 Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, rosids
 / eurosids II, Brassicales; Brassicaceae; Arabidopsis.
AUTHORS Pih,K.T., Park,J.M., Jang,H.J., Kang,S.G., Piao,H.L. and Hwang,I.
TITLE EST of salt inducible mRNA in Arabidopsis thaliana
JOURNAL Unpublished
COMMENT Contact: Kyeong Tae Pih
 Department of Plant Molecular Biology

Gyeongsang National Univ., Plant Molecular Biology and
 Biotechnology Research Center
 Jinju, Kyungnam 660-701, Korea
 Tel: 82-591-751-5193
 Fax: 82-591-759-9363
 Email: ihwang@ongae.gsnu.ac.kr
 Seq primer: 73
 High quality sequence scop: 307.

FEATURES
source
 1..307
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /dev_stage="7 day-old"
 /clone_lib="NaCl-treated Arabidopsis subtraction library"
 /note="Subtracted cDNA library from salt (NaCl)-treated whole plants"

BASE COUNT 71 a 197 c 41 g 96 t 2 others
ORIGIN

Query Match 10.7%; Score 288.8; DB 9; Length 307;
Best Local Similarity 97.0%; Pred. No. 5e-3e;
Matches 293; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1792 CTTTATGAGCGTTTCAAGTCCATCCATCTTACCAAGTTGCGCGTCTTTCATCA 1851
 DB 1 CTTTATGAGCGTTTCAAGTCCATCCATCTTACCAAGTTGCGCGTCTTTCATCA 60
 QY 1852 TTCCGAGTCTTCTGCGGTGAACCAATACAGATGCTCTCAATTCCTTC 1911
 DB 61 TTCCGAGTCTTCTGCGGTGAACCAATACAGATGCTCTCAATTCCTTC 120
 QY 1912 CTCAGAAATGGTGGCATTCCTTCTTCTTCCCTCAAGATTCGAAAGTGTCCA 1971
 DB 121 CTCAGAAATGGTGGCATTCCTTCTTCTTCCCTCAAGATTCGAAAGTGTCCA 180
 QY 1972 GTTACCCACCGGTCTCATGCGGATCTTAAGACCGGAACCGTGGCTTAATGCGGA 2031
 DB 181 GTTACCCACCGGTCTCATGCGGATCTTAAGACCGGAACCGTGGCTTAATGCGGA 240
 QY 2032 TCTATGCTAAGCATCATGCGCATTCGTAACATCAATGCGGCTTTAATATTC 2091
 DB 241 TCTATGCTAAGCATCATGCGCATTCGTAACATCAATGCGGCTTTAATATTC 300
 QY 2092 TT 2093
 DB 301 TT 302

RESULT 15 602 bp DNA linear GSS 13-DEC-2002
B2476824
LOCUS BONB01TR BQ.1.6.2 KB loc Brassica oleracea genomic clone BONB01,
DEFINITION genomic survey sequence.
ACCESSION B2476824
VERSION B2476824.1 GI:26779222
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE Eukaryote, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;
 Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, rosids
 / eurosids II, Brassicales; Brassicaceae; Brassica.
AUTHORS Town,C.D., Van Aken,S., Utecherback,T., Koo,H. and Frazer,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished
COMMENT Other_GSSs: BONB01TF
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-8523
 Fax: 301-838-0208

Email: cdt@owenlgr.org
DNA is from a doubled haploid provided by Tom Osborn
Seq primer: TR
Class: sheared ends.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| Source | 1. .602 |

| | | | | |
|------------|-------|-------|-------|-------|
| BASE COUNT | 168 a | 109 c | 198 g | 127 t |
| ORIGIN | | | | |

| | | | | |
|-----------------------|-----------------|--------------------|------------|-------------|
| Query Match | 10.4% | Score 280; | DB 29; | Length 602; |
| Best Local Similarity | 74.1%; | Pred. No. 1.1e-34; | | |
| Matches 464; | Conservative 0; | Mismatches 120; | Indels 42; | Gaps 7 |

| | | | |
|----|------|--|------|
| OY | 1724 | CTTTACCTGGTAGAGATAGCTGGGACCTTCAATCTCGAGTAATCTCAATA---TTGATCAAGCA | 1780 |
| Db | 602 | CTCTGCTGGTAGACATAGCTGGATCATCTCTCGAGCCACTCTTAACTTTCTTTATATA | 543 |
| OY | 1781 | ACCAAGCCCTCTTTCATCAGCCCTTTCAAAGTCCCATCCCATCTTACCAGATTAGCCCGT | 1840 |
| Db | 542 | ACCAAGCCCT-----TTGAGAGATGCCATCTTCTTACCAAGTCAATGCCAT | 495 |
| OY | 1841 | CTTCTTATCATATCCCGAGTCTCTCTCCGGGTGGAAACCAATTAACAATATGCTCTTACAT | 1900 |
| Db | 494 | CGTCTTCTTCAATCCCTAGCCCTTCTTCGGG-----GCACAACATCTCACTA | 447 |
| OY | 1901 | TCTTCCCTTCTCCAGAAATGGTGGCAATTCCTTCTTCTTCTTCCCTCCCAATCTCGAA | 1960 |
| Db | 446 | TCTTCCCTTCTCCAGAAACGGTGGGATCCCTTCTGCTGCTCTCTCTCAAGATCTCAA | 387 |
| OY | 1961 | ACAATTTGTCAGATTACCCCAACGGGTCTCATGGCCCACTTCTAAGAACCCGAAACCGTTGC | 2020 |
| Db | 386 | ACAGTGTCTCCCTCACTCCACTCGTCTCTTCCCAACATTAATMAACCCCAACCGGTTAC | 327 |
| OY | 2021 | CTAATCTGGGAATCTATCGCTAAGCAATCCATGAGGCAT---TGTCTAAACAATGATGGGT | 2077 |
| Db | 326 | CTACTTGGGAGCTTTTACCAACAGTGCATGGCCATTGGCTGTAAACAGTCAATGTCCTT | 267 |
| OY | 2078 | CTTTTAATTAATCTTTCTATGCGGATTTCTGCACTGTAGTCCGACAGATGCCACAGT | 2137 |
| Db | 266 | CTTTTAATGTAACCGGTTCTTACGGGGTTTCTGCGCCGCGAGTCCCACTATATGCCAGT | 207 |
| OY | 2138 | TTCAATACCCCTGGCTACTATACTGTAATGTGATGAGTCTGACTCTTCCACTGTGATTCTG | 2197 |
| Db | 206 | TCAATGCAACCGGCTACTATACTCTGAGTGCAGAGATCTGAGCCCTTCAACTGTGATTCTG | 147 |
| OY | 2198 | GTCATTTGGATAAGCTTTGAGAAAGTTGGACAA---CAACAGGCATTCTCTG-----CCT | 2248 |
| Db | 146 | GTCATTTGGATAAGCTTTGAGAAAGTTCTCAACAACAGCAACAGCCGTTCCATGGGGTGTCT | 87 |
| OY | 2249 | CTATGTGTCCAAACCTCTCCTAACCCTTCAATCTTGTGA---ACCTGCGCCTCAGCAGATGT | 2305 |
| Db | 86 | CTGCAAGTCCCGGCTTCTCTCACTTCAACCTAAGTAACACAGATGCTCAGGGGGCTAT | 27 |
| OY | 2306 | CTCCAAATATAGTGCCTTCCCAAGG | 2331 |
| Db | 26 | CGGCAACACATGCGGGAATCCCAAGG | 1 |

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 19:12:39 ; Search time 153 Seconds
(without alignments)
7751.611 Million cell updates/sec

Title: US-09-995-938a-2

Perfect score: 2687

Sequence: 1 accatgcttcgcctcaatc.....attattctgagataaaaaa 2687

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/CTUS_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|---------|-------|--------------------------------------|
| 1 | 66.6 | 2.5 | 7218 | 1 | US-08-232-463-14 Sequence 14, Appl |
| 2 | 50.4 | 1.9 | 12730 | 2 | US-09-004-838-91 Sequence 91, Appl |
| 3 | 48.4 | 1.8 | 19124 | 2 | US-08-487-8268-13 Sequence 13, Appl |
| 4 | 47.8 | 1.8 | 15062 | 4 | US-09-004-838-89 Sequence 89, Appl |
| 5 | 46.6 | 1.7 | 2435 | 3 | US-09-306-593-1 Sequence 1, Appl |
| 6 | 46.2 | 1.7 | 7218 | 1 | US-08-332-463-14 Sequence 14, Appl |
| 7 | 46.2 | 1.7 | 9510 | 4 | US-09-453-7028-256 Sequence 256, App |
| 8 | 45.2 | 1.7 | 1511 | 1 | US-07-991-8678-8 Sequence 8, Appl |
| 9 | 45.2 | 1.7 | 1511 | 2 | US-08-107-7554-8 Sequence 8, Appl |
| 10 | 45.2 | 1.7 | 1511 | 2 | US-08-544-332-8 Sequence 8, Appl |
| 11 | 45.2 | 1.7 | 1511 | 4 | US-09-370-8614-8 Sequence 8, Appl |
| 12 | 44.6 | 1.7 | 168575 | 4 | US-09-426-290-1 Sequence 8, Appl |
| 13 | 44.4 | 1.7 | 12793 | 4 | US-09-004-838-124 Sequence 124, App |
| 14 | 43.4 | 1.6 | 636 | 3 | US-08-998-416-1137 Sequence 1137, Ap |
| 15 | 43.4 | 1.6 | 14507 | 3 | US-08-785-150-1 Sequence 1, Appl |
| 16 | 43.4 | 1.6 | 14507 | 4 | US-09-660-299-1 Sequence 1, Appl |
| 17 | 43.4 | 1.6 | 14507 | 4 | US-09-435-377-1 Sequence 1, Appl |
| 18 | 43.2 | 1.6 | 615 | 3 | US-08-998-416-186 Sequence 186, App |
| 19 | 43.2 | 1.6 | 4262 | 4 | US-09-521-511C-10 Sequence 10, Appl |
| 20 | 42.8 | 1.6 | 3715 | 4 | US-09-041-886-10 Sequence 10, Appl |
| 21 | 42.8 | 1.6 | 1664976 | 4 | US-08-916-421B-1 Sequence 1, Appl |
| 22 | 42.6 | 1.6 | 1308 | 4 | US-09-601-198-3 Sequence 3, Appl |
| 23 | 42.2 | 1.6 | 53332 | 4 | US-09-801-861-3 Sequence 3, Appl |
| 24 | 42 | 1.6 | 854 | 3 | US-08-998-416-534 Sequence 534, App |
| 25 | 41.8 | 1.6 | 5852 | 1 | US-07-867-106-2 Sequence 2, Appl |
| 26 | 41.6 | 1.5 | 860 | 3 | US-08-998-416-287 Sequence 287, App |
| 27 | 41.4 | 1.5 | 15016 | 4 | US-09-601-198-60 Sequence 60, Appl |

| | | | | | |
|----|------|-----|--------|---|--------------------------------------|
| 28 | 41.4 | 1.5 | 168575 | 4 | US-09-426-290-1 Sequence 1, Appl |
| 29 | 41.2 | 1.5 | 240 | 1 | US-08-628-417-6 Sequence 6, Appl |
| 30 | 41.2 | 1.5 | 732 | 3 | US-08-998-416-1036 Sequence 1036, Ap |
| 31 | 41.2 | 1.5 | 1798 | 4 | US-09-797-906-1 Sequence 1, Appl |
| 32 | 41.2 | 1.5 | 19124 | 2 | US-08-487-8268-13 Sequence 13, Appl |
| 33 | 41.2 | 1.5 | 19124 | 2 | US-08-487-8268-13 Sequence 7, Appl |
| 34 | 41 | 1.5 | 5852 | 1 | US-07-867-106-2 Sequence 2, Appl |
| 35 | 40.8 | 1.5 | 658 | 3 | US-08-998-416-595 Sequence 595, App |
| 36 | 40.8 | 1.5 | 1496 | 4 | US-09-835-811-1 Sequence 1, Appl |
| 37 | 40.6 | 1.5 | 4673 | 1 | US-07-638-431-1 Sequence 1, Appl |
| 38 | 40.6 | 1.5 | 4673 | 5 | PCT-US92-00018-X Sequence 1, Appl |
| 39 | 40.6 | 1.5 | 6243 | 2 | US-09-056-075-1 Sequence 1, Appl |
| 40 | 40.4 | 1.5 | 7832 | 4 | US-09-004-838-94 Sequence 94, Appl |
| 41 | 40.4 | 1.5 | 10007 | 4 | US-09-417-485D-13 Sequence 13, Appl |
| 42 | 40.4 | 1.5 | 10040 | 4 | US-09-417-485D-13 Sequence 5, Appl |
| 43 | 40.2 | 1.5 | 658 | 3 | US-08-998-416-595 Sequence 595, App |
| 44 | 40.2 | 1.5 | 872 | 3 | US-08-998-416-407 Sequence 487, App |
| 45 | 40.2 | 1.5 | 6124 | 4 | US-08-213-419B-3 Sequence 3, Appl |

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER TYPE: FLOPPY disk
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 INMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 846-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZpT-Fls
; US-08-232-463-14
Query Match 2.5%; Score 66.6; DB 1; Length 7218;

Best Local Similarity 3.8%; Pred. No. 2,1e-07;
Matches 15; Conservative 233; Mismatches 147; Indels 0; Gaps 0;

QY 1793 TTTCATCAGCCTTTCAGAGTCCCATCCATCCATCCAGTACGCGCTTCTTCATCAT 1852
1067 YY 1126
QY 1853 TCCCGAGTCTTCTCGGGGGAACCAATAACAACATGCTCTACATTTCTCCCTTCC 1912
1127 YY 1186
QY 1913 TCAGAAATGGTGGCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1972
1187 YY 1246
QY 1973 TTAACCCACGCGTCTCATCGCGCATCTTACAGAACCGGAAACCGTCTCTACTGGAA 2032
1247 YY 1306
QY 2033 CTATCGTAAAGCAATCCATGCGCATCTCTAAACATCAATGCGCTTTTAATTATCCTT 2092
1307 YY 1366
QY 2093 TCTATGCGCTTTCGACGCTGCTAGTCCGACATGCGCAACGATTCTCATACCGTGGCTA 2152
1367 YY 1426
QY 2153 CTATACCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2187
1427 YYYYYYYYGTACCAAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1461
Db

RESULT 2

US-09-004-838-91/c
Sequence 91, Application US/09004838
Patent No. 6350933
GENERAL INFORMATION:
APPLICANT: Michelmore, Richard W.
APPLICANT: Shen, Kathy
APPLICANT: Meyers, Blake
TITLE OF INVENTION: Procedures and Materials for
TITLE OF INVENTION: Confering Pest Resistance in Plants
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,838
FILING DATE: 09-JAN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/781,734
FILING DATE: 10-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Einhorn, Gregory P.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 023070-078810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 12730 base pairs
TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: -

LOCATION: 1..12730

OTHER INFORMATION: /note= "RG2C"

US-09-004-838-91

Query Match

Best Local Similarity 47.9%; Pred. No. 0.0059;

Matches 205; Conservative 0; Mismatches 221; Indels 2; Gaps 2;

QY 173 GCCGAACCTTGTATGAGCAAACTGCTAAACCCCTTTCTTTATGTTCAATGAAC 232
7038 GGCAGAAACGTAAAT 6979
QY 233 ATACAGTTTGTGTTATGATACATAATATATATATATATATATATATATATATATAT 292
6978 TATATATGTAATGTAAGAGGAAAGTGTCAAAATATATATATATATATATATATATAT 6919
QY 293 GATATATAGATCTAAAGCTTAAATATCAATGTAACCAACTAAATATATATATATAT 351
6918 TGAATATATAGATCTTAT 6859
QY 352 ATATGTAATGTAAT 411
6858 TTCTTAAAGCTTATGTAAT 6799
QY 412 AGAGAGAGCGATGTACATCAACTATATATATATATATATATATATATATATATATAT 471
6798 TAAAGAACTAAATTTGAAGCAACTGAAATTTTAAATAA-TTAAATTCATGTAATTTAA 6740
QY 472 TGAGATTGTATGCAAGAAAGTGAATTTTAAATGTAATATATATATATATATATATAT 531
6739 ATTAATTTTAT 6680
QY 532 ATCAACGCTTCACTGAT 591
6679 GCAT 6620
Db

RESULT 3

US-08-487-8268-13
Sequence 13, Application US/084878268
Patent No. 593827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chien, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knodde Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NH121.001CPL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 1.8%; Score 48.4; DB 2; Length 19124;
Best Local Similarity 50.6%; Pred. No. 0.025;
Matches 170; Conservative 0; Mismatches 161; Indels 5; Gaps 2;

QY 246 TTATGATTCATTAATTAATGATGACCCGACCAATTAATCCAAATTTGATATTGATATG 305
DB 15614 TTTTAT 15673
QY 306 TAAAGCTTAAATCAACATGTAACCAATTAATTAATTAATTAATTAATTAATTAATTA 365
DB 15674 ATTAAT 15732
QY 366 TCACCAATCTTAT 425
DB 15723 --ATAAAT 15790
QY 426 TACATCACTATATCAAGCAAAAGTCAAAATTTCTTAAATGATGATTTGATATG 485
DB 15791 TACATCACTATATCAAGCAAAAGTCAAAATTTCTTAAATGATGATTTGATATG 15850
QY 486 AAGAAAGAGTAT 543
DB 15861 AAAT 15910
QY 544 CACTGAT 579
DB 15911 AATGAAAT 15946

RESULT 4

US-09-004-838-89/c

Sequence 89, Application US/09004838

Patent No. 6350933

GENERAL INFORMATION:

APPLICANT: Michelmore, Richard W.

APPLICANT: Shen, Kathy

APPLICANT: Meyers, Blake

TITLE OF INVENTION: Procedures and Materials for

NUMBER OF SEQUENCES: 140

CORRESPONDENCE ADDRESSES:

ADDRESSES: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/004,838

FILING DATE: 09-JAN-1998

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/781,734

FILING DATE: 10-JAN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Einhorn, Gregory P.

REGISTRATION NUMBER: 38,440

REFERENCE/DOCKET NUMBER: 023070-078810US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 89:

SEQUENCE CHARACTERISTICS:

LENGTH: 15062 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: 1.15062

LOCATION: 1.15062

OTHER INFORMATION: /note="RG2B"

US-09-004-838-89

Query Match 1.8%; Score 47.8; DB 4; Length 15062;
Best Local Similarity 47.4%; Pred. No. 0.032;
Matches 203; Conservative 0; Mismatches 223; Indels 2; Gaps 2;

QY 173 GCCGAACCTGTTATGGAACAACCTGTAACCCCTTTTCTTATGTTCAATGACT 232
DB 9453 GGCAGAACCTGTAAT 9394
QY 233 ATACAGTTTGTGATGATATATATATATATATATATATATATATATATATATAT 292
DB 9393 TAT 9335
QY 293 GGATATGATATCTAAAGCTTAATATATATATATATATATATATATATATATAT 352
DB 9334 GTTAGTATATCTGTAATATATATATATATATATATATATATATATATATAT 9275
QY 353 TATGTAATGAT 411
DB 9274 ATGCAAT 9215
QY 412 AAGAGAGCCAGTAT 471
DB 9214 TTAAGAACTCAATTTGAAGCAAACTGAAATTTTAAATATATATATATATAT 9155
QY 472 TGAGATTGTATGCAAGAAAGTATATATATATATATATATATATATATATATAT 531
DB 9154 ATTAATTTACATTTTCAATATATATATATATATATATATATATATATATAT 9095
QY 532 ATCAAGGCTTCACTGTAATATATATATATATATATATATATATATATATAT 591
DB 9094 GCAAT 9035
QY 592 GGTCAAC 599
DB 9034 TATATATG 9027

RESULT 5

US-09-306-593-1

Sequence 1, Application US/09306593

Patent No. 6184018

GENERAL INFORMATION:

APPLICANT: Li, Xin-Liang

APPLICANT: Ljungdahl, Lars G.

APPLICANT: Chen, Huizhong

APPLICANT: Kimes, Eduardo A.

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296,95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 256:
SEQUENCE CHARACTERISTICS:
LENGTH: 9510 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 256:
US-09-453-702B-256

Query Match 1.7%; Score 46.2; DB 4; Length 9510;
Best Local Similarity 51.2%; Pred. No. 0.066;
Matches 108; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Oy 276 CAATTAATCCAAATTTGGATTGATGATCTAAAGTTAAATCAACATGTAAACCAACT 335
|||
Db 1185 CAACATTTTACATTTGATGATCAATGAAATATCAATGATGATGATGATGATGAT 1244
|||
Oy 336 AAATCTTATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 395
|||
Db 1245 TGACTTTTAACTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTT 1304
|||
Oy 336 GAAGAGAGTAAAG 455
|||
Db 1305 AGTAG 1364
|||
Oy 456 AAAATTGTTCTAAAGTGAATTTGATGCA 486
|||
Db 1365 GGATAGATTCAACTGAGAGATTTAAACA 1395
|||

RESULT 8
US-07-991-867B-8
Sequence 8, Application US/07991867B
Patent No. 5476781
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidi, Michael E.
TITLE OF INVENTION: No. 5476781 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/991,867B
FILING DATE: 12-DEC-1992
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UP14.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amsacta moorei entomopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511

US-07-991-867B-8

Query Match 1.7%; Score 45.2; DB 1; Length 1511;
Best Local Similarity 44.7%; Pred. No. 0.043;
Matches 176; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

Oy 181 TTGTTATGGAGAACTGGTAAACCCCTTTTCCTTTATGTTCAATGACATGATAGT 240
|||
Db 737 TTGTTTGGCAAGAAACATAGAGACCAATTTATTAATTCATGACATTTTATTATT 796
|||
Oy 241 TTGTTATGATATCATTAATTAATGATGAGACCAACCAATTAATCCAAATTTGAT 300
|||
Db 797 TGATATATTTTTCACAAAAAATTAATCATGAGAAAAAATTAATTAATCAAAATGA 856
|||
Oy 301 GATACATTAAGCTTAATTAATCAATGATGATGATGATGATGATGATGATGATG 360
|||
Db 857 TTTACTAAATTTGATATATTTTAAATTAATTTTAAATTAATTTAATTTAATTTA 916
|||
Oy 361 GGTATTCACCAATCTTATATCATTTGTAAGTACAGAGAGTAAAGAGAGAGAG 420
|||
Db 917 AATTAATTAACAGAGAAATGTTATTAATTAATTAATTAATTAATTAATTAATTA 976
|||
Oy 421 CAGTGTACATCAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
|||
Db 977 AGAATTCAT 1036
|||
Oy 481 TAGCAAG 540
|||
Db 1037 AAGTTATATATTTCAAAATTTAATTAATTTAATTTAATTTAATTTAATTTA 1096
|||
Oy 541 CTCACGTATTAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 574
|||
Db 1097 TATTAATATTTGATATATATTTAATTAATTAATTAATTAATTAATTAATTA 1130
|||

RESULT 9
US-08-107-755A-8
Sequence 8, Application US/08107755A
Patent No. 5721352

GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5721352e1 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/107,755A
FILING DATE: 19-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,658
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF114.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Ambacta moorei entemopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-08-107-755A-8

Query Match 1.7%; Score 45.2; DB 1; Length 1511;
Best Local Similarity 44.7%; Pred. No. 0.043;
Matches 176; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

181 TTGTTATGAGCAAACTCGTAAACCCCTTTTCCTTTATGTTCAATGAACTATACAGT 240
|||||
737 TTGTTTTCGACAAAACATGAGCCAAATTAATCTATCGCATTTTATTTATTTAT 796
|||||
241 TTGCGTATGATACATAAATAATGATGAGCCAGCAATTAATCCAAAATTGGATATTA 300
|||||
797 TGATATATTTTTCAAAATAATTAATCAATGAAAAAATAAATAATATCAAAATGGA 856
|||||
301 GATACCTAAGCTTAAATCAACATGTAACCAACTAATCTTAATAGAACTAGTAAT 360
|||||
857 TTCTACAAATTTGATATATTTTAAATTAATATTTAAATATTAATTTAAAAAAT 916
|||||
361 GGTATTCACCAATCTTATATCATTTTGTAGTACGAGAGGTAAAAAAGAGAGACG 420
|||||

Db 917 AATAATAAACAGATATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 976
Qy 421 CAGTGACATACACTAATACAGACAAAAGTACTCAAAATGTTTCTTAAAGTGAATTTG 480
|||||
Db 977 AGAATGCAATTA 1036
|||||
Qy 481 TATCAAGAAAAAGATGAATTTTAAATGATATATCATTTTNGATTTAATCAACGCG 540
|||||
Db 1037 AAGTTATATATTTCAAAATTTAAATATTAATTAATTTTAAATTTTAAATTAACAA 1096
|||||
Qy 541 CTTCACTGTAATAAATTAATTTTAAACCAAT 574
|||||
Db 1097 TATACATATTTAGATATATCTTATACAAAAAT 1130
|||||

RESULT 10
US-08-544-332-8
Sequence 8, Application US/08544332
Patent No. 5935777
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5935777e1 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gerard H. Bencen
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,867
FILING DATE: 07-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,755
FILING DATE: 19-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerald H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UF114.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Ambacta moorei entemopoxvirus
FEATURE:

TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1137:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1692RP
US-08-998-416-1137

Query Match 1.6%; Score 43.4; DB 3; Length 636;
Best Local Similarity 47.0%; Pred. No. 0.079;
Matches 166; Conservative 0; Mismatches 186; Indels 1; Gaps 1;

QY 226 ATGACTATACAGTTTGGTATGATCATTAATTAATGATGAGCCAGCAATTAATCC 285
DB 571 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 512
QY 286 AAAATTTGGATATTAGATCTTAAGCTTAATCAACATGTAACCAACTAATACCTTA 345
DB 511 AATAATATAAATTAATTAATTAAGAAATTAAGTAAATTAATTAATTAATTAATTCCT 452
QY 346 TAGAATCATGTAAATGATTCACCAATCTTTATATCATTTGTAAGTACGAAGAAGTA 405
DB 451 ATTAATAAGATTAATTAATTAATTAATCAATCAATTAATTAATTAATTAATTA 393
QY 406 AAAAAAGAGAGAGCCAGTACATCACTATCAAGCAAAAGTAGCAAAATGTTT 465
DB 392 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 333
QY 466 CTAAAGTGAATTTGATGCAAGAAAAAGTGAATTTTAAATGATATATCATTTATG 525
DB 332 TAATAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 273
QY 526 ATGTTAATCAGACGCTTCACTGTATTAATTAATTAATTAATTAATTAATTAATTA 578
DB 272 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 220

RESULT 15
US-08-785-150-1/c
Sequence 1, Application US/08785150
Patent No. 6027915

GENERAL INFORMATION:
APPLICANT: Moritz, Arvia E.
APPLICANT: Lee, Chi-Chang
APPLICANT: Thomas, James N.
TITLE OF INVENTION: Expression Augmenting Sequence Elements
Patent No. 6027915
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Operating System Software 7.1
SOFTWARE: Microsoft Word for Macintosh, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785,150
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/586,509
FILING DATE: 11-JAN-96
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2841
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14507 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: NO. 6027915 Relevant
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Chinese hamster
IMMEDIATE SOURCE:
CLONE: 2A5-3 lambda CHO sequence
US-08-785-150-1

Query Match 1.6%; Score 43.4; DB 3; Length 14507;
Best Local Similarity 48.9%; Pred. No. 0.47;
Matches 116; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 469 AAGTGAATTTGATGCAAGAAAAAGTGAATTTTAAATGATATATCATTTATGATG 528
DB 13710 AAAAGGATTAAGTGAATTAAGAAATTTGACCAACCAATGATTAATTAATTA 13651
QY 529 TTAATCAGACGCTTCACTGTATTAATTAATTAATTAATTAATTAATTAATTA 588
DB 13650 TTAAGCAGACGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 13591
QY 589 TTGCTGACCAAGTAAGGAGCCACACTGAAGAACGCTGCCACTGTCTCTCTTT 648
DB 13590 TTCTGAAAAAATTAAGAGCAAAATTAATTAATTAATTAATTAATTAATTA 13531
QY 649 CTTTCTCTGATTAATTTGGTCAATCTCAATTTTACCTACAGAAAAAATTA 705
DB 13530 CTTCAAAAGATGCTATTCTTAATTAATTAATTAATTAATTAATTAATTAATTA 13474

Search completed: December 19, 2003, 20:59:43
Job time: 155 secs